Sensitive and Specific Proteomic Identification and Quantitation of Murine Cytochrome P450 Enzymes and Histone Post-Translational Modifications using Mass Spectrometry

by
Elisabeth M. Hersman

A dissertation submitted to the Johns Hopkins University in conformity with requirements for the degree of Doctor of Philosophy

Baltimore, MD
December 2013

Abstract

Mouse models are widely used in pharmacology, yet the expression profile of the murine drug metabolizing enzymes has only begun to be characterized at the protein level. We developed a quantitative, high-throughput mass spectrometric method to measure the protein expression of the cytochromes P450 (Cyps) in tissue lysates isolated from Balb/c mice. Global mass spectrometry-based proteomics revealed 27 proteins belonging to Cyp subfamilies 1a, 2a, 2b, 2c, 2d, 2e, 2f, 2j, 2u, 3a, 4a, 4b, 4f, and 4v were readily detectable in Balb/c mouse tissue. Using this protein list, a selected reaction monitoring mass spectrometric screen was developed to quantify expression of these 27 proteins. The screen was applied to mouse liver microsomes and tissue lysates of kidney, lung, intestine, heart and brain from mixed sex fetuses; 3-4 weeks, 9-10 weeks, and 8-10 months of age male and female mice; and pregnant mice. All 27 proteins were identified in liver microsomes. Cyps were identified in all lysates, including 10 in kidney, 15 in lung, 8 in intestine, 6 in heart, and 4 in brain. Liver microsomes differed in expression with sex, age, and pregnancy. Cyps 1a2, 2c67, and 4a12a appeared to be more abundant from male mouse livers of all ages. Hepatic expression of Cyp2b9 was more abundant in 3-4 week old mice in comparison to mice of other ages, and it was the only enzyme identified in higher abundance in pregnant mouse microsomes in comparison to the age matched females. Sexually dimorphic expression appeared exclusively in kidney for Cyps 2b9, 2d26, 2e1, and 4b1. Microsome activity experiments evaluated the mouse liver microsome samples for Cyp activity against the HIV maintenance drug,

efavirenz, revealing differences in metabolite formation with age, sex, and pregnancy status. Collectively, these data provide a foundation for the use of mice as a preclinical model for understanding Cyp pharmacology and contributes to a wider understanding of mouse xenobiotic metabolism.

In a second set of experiments, a derivatization technique of unmodified lysine residues on histones is used in combination with high performance mass spectrometry to distinguish and quantitate endogenously acetylated isoforms occurring within the same tryptic peptide sequence and to extend this derivatization strategy to other post-translational modifications, specifically methylation, dimethylation and trimethylation. The in vitro deuteroacetylation of monomethylated lysine residues is observed, though dimethylated or trimethylated residues are not derivatized. Comparison of the relative intensities ascribed to the deuteroacetylated and monomethylated species with the deuteroacetylated but unmethylated analog, provides an opportunity to estimate the percent of methylation at that site. In addition to the observed fragmentation patterns, the very high mass accuracy available on the Orbitrap mass spectrometer can be used to confirm the structural isoforms, and in particular to distinguish between trimethylated and acetylated species.

Readers: Namandjé Bumpus, Katherine Wilson

Acknowledgements

My work at the Hopkins School of Medicine has been made possible by a few phenomenal individuals and the incredibly supportive Pharmacology Department community. I thank Kathy Wilson, Jef Boeke, Phil Cole, Stefani Thomas, Lerna Uzasci, Christine Jelinek, and Julie Lade for their mentorship and/or professional support.

Bob Cotter, who passed away in November 2012, was a light-hearted boss who always assumed his students to be proficient at the next level, earning him the endearing name “Whiz-Kid”. He allowed his students the freedom to create projects and to become scientists. His quiet appreciation for every individual and their own stage of learning is what I most admired and will aspire to emulate.

Namandjé Bumpus, not knowing who I was except from her interactions with Bob, offered me her guidance and support after Bob passed away. I was in awe of her commitment to our community, which I continued to see throughout the next 8 months working with her. I don’t think I realize yet all the ways she has inspired me, but I know I will be eternally grateful of her mentorship.

Lastly, I thank my family, whose love, encouragement, and high standards keep me grounded. There are no words to describe my appreciation for my partner Dave Walker, who is unfailingly patient and giving in all aspects of his life. Dave’s generous personal, professional, and technical support has been essential to my success at Hopkins.

Table of Contents

Thesis paper for graduate research in the labs of

Namandjé N. Bumpus and Robert J. Cotter

List of Tables

Page

List of Figures

Page

Abbreviations

Collision induced dissociation (CID)

Cytochrome P450, human (CYP)

Cytochrome P450, mouse (Cyp)

Electrospray ionization (ESI)

High performance liquid chromatography (HPLC)

Mass to charge ratio (*m/z*)

Matrix Associated Laser Desorption Ionization (MALDI)

Multiple Reaction Monitoring (MRM)

National Center for Biotechnology Information Reference Sequence Database (NCBI RefSeq Database)

NCBI Basic Local Alignment Search Tool (BLAST)

Nicotinamide (NIA)

Peptide fragment, C-terminal (y-ion)

Peptide fragment, N-terminal (b-ion)

Post-translational modification (PTM)

Selected Reaction Monitoring (SRM)

Time of flight mass spectrometer (TOF)

Trichostatin A (TSA)

Chapter 1:

Introduction

Proteomics

"Proteomics" is the global study of cellular proteins. Of the four types of different macromolecules, including proteins, lipids, carbohydrates and nucleic acids, proteins are the primary mediators of cellular activity (Tyers & Mann, 2003). Proteins are involved in structural organization, signaling between cellular compartments, and enzymatic reactions to chemically change another molecule. They work together through associations and conformational changes and can be regulated by protein-mediated transcription/translation, degradation, or modification (Tyers & Mann, 2003). Probing cellular dynamics as a whole protein network offers a more complete view of cellular interactions and activity (Clancy & Hovig, 2013; Tian, 2013). Currently, many "-omics" fields are devoted to global research using high-throughput methodology, a trend which began with "genomics" in the 1980's (Kuska, 1998). Once the genome was published, theoretical proteins could be transcribed electronically "in silico." Global discovery of proteins using mass spectrometry is now well-established using the genome databases, and now the focus is on the direction of biology (Ahmad & Lamond, 2013). Researchers are applying proteomics to technically challenging projects such as the quantification of highly modified proteins or highly similar proteins.

Proteins are globular molecules made from multimers of amino acids. All 20 amino acids have the formula R-CH(NH2)COOH, where the amine and the carbonyl group form a polar protein backbone. The R group can be acidic, basic,

nonpolar, small (56 Da) or large (186 Da), each denoted by a single letter for simplicity in recording their sequence. The sequence is listed from the amino-terminus (Waterston et al) to the carboxy terminus (C-). Some R-groups, for example lysine (K) and arginine (R), can be modified in the cell to receive a "post-translational modification", for example acetylation ($-COCH_3$) or methylation ($-CH_3$). The protein is composed of a sequence of residues, which reflects the function of the protein. For example, the residues exposed to cytosol are typically hydrophilic and the residues in the protein core are more hydrophobic (Jager et al, 2008). Histones have a hydrophilic tail containing many lysine and arginine residues which are available to cellular enzymes for modification in addition to a hydrophobic core. In contrast, cytochrome CYP enzymes are embedded in the endoplasmic reticulum membrane and their function requires nonpolar substrates to diffuse into the active site, which is reflected in their sequence. The type and sequence of residues in a protein are unique, which underlies the power of proteomics and contributes to its technical challenges. In a typical proteomics analysis, proteases digest the protein into peptides an appropriate size for determining the sequence. The most commonly used protease is trypsin, which cleaves the amide peptide bond on the carboxyl side of the basic amino acid residues lysine and arginine (Olsen et al, 2004). Proteolysis by trypsin typically yields peptides ranging from 200-4000 Da, which are small enough to be detected by high resolution mass spectrometry and which contain enough residues for a unique sequence.

Mass Spectrometers

Mass spectrometry identifies the masses of molecules in a sample. The instrument induces a charge on the molecule by gain or loss of a hydrogen ion, and the subsequent ions are controlled by electrostatic forces in the instrument. As ions move towards an opposite charge, the acceleration is a function of the mass to charge ratio (*m/z*). Larger molecules change their movement more slowly than smaller molecules of the same charge, and in this way, the movement of the ions is indicative of the mass. Mass spectrometers usually include a mechanism for isolating and breaking an analyte ("parent compound") into multiple fragments ("transitions"). The experiments in this thesis use a fragmentation method called collision induced dissociation (CID) whereby an inert gas such as nitrogen or helium is introduced into the vacuum. Multiple collisions impart enough vibrational energy within the peptide to break the backbone. The half of the fragment that receives the charge is subsequently controlled by the optics of the instrument for later detection. The fragment can either be the N-terminal fragment (b-ion) or the C-terminal fragment (y-ion). The process of collecting a mass spectrum from the parent and then from a fragment is called MS/MS. The MS/MS spectrum is considered to be the fingerprint for identification of the analyte. Instruments differ in the particular combination of method of ionization, ion separation, and detection systems. The time of flight, Orbitrap, and triple quadrupole used in this research are discussed below.

The time of flight (TOF) instrument detects *m/z* based on separation in a vacuum and the arrival time of the ions at the detector (Cotter & American Chemical Society. Meeting, 1994). The TOF imparts an extraction potential at the site of ionization so that oppositely charged ions accelerate toward it. The potential applies the same amount of force on all ions, but the larger molecules accelerate more gradually and therefore travel with less velocity. Ions continue their momentum down the flight tube. A detector records the abundance of ions as a function of time with the time of flight as the indicator of mass. Sensitivity can be as low as a single molecule, depending on the TOF instrument. The distribution of ions is disproportional such that molecules with low molecular weight have high velocities and overlapping arrival times whereas molecules of higher molecular weights are readily distinguished. Many TOF mass spectrometers incorporate a curved field reflectron to accomplish four-fold higher resolution (20,000 FWHM) while maintaining a high mass range of 80 kDa (ShimadzuBiotech, 2006). This technique is often paired with a solid sample ionization method called Matrix Associated Laser Desorption Ionization (MALDI). It employs a matrix such as alpha-cyano-4-hydroxycinnamic acid to stabilize the sample and to impart a charge. The dried sample-matrix spot is shot with a soft 337nm laser, which imparts enough energy to vaporize the analyte for detection in the mass spectrometer. Molecules of a known weight are used to calibrate the time of flight to the *m/z*. Resolution and accuracy are most affected by variable travel distance to the end of the extraction region. While a thick sample

deposition or a large distance between the positions of the reference sample to the experimental sample on the sample plate can noticeably alter resolution and accuracy, this method is nevertheless popular for analysis of a relatively simple mixture, especially for labile modifications and larger peptides and proteins.

In 2000, the invention of the Orbitrap allowed high resolution analysis of many ions simultaneously (Makarov, 2000). The Orbitrap is composed of a metal rod, a nearly seamless metal casing, and a vacuum area between all about 1.5 inches wide and 1 inch tall. Ions are injected into the Orbitrap in a packet. The ions enter the Orbitrap in a trajectory to fly past the rod, but the charged rod attracts the ions such that they go into a stable orbit. The rod is thinner on the ends than in the middle of the chamber. This shape forces ions that reach the end to move back towards the middle. As the ions are orbiting they have a characteristic frequency of movement from one side of the rod to the other. The movement of ions in orbit creates an electromagnetic signal. The total signal strength as a function of time can be deconvoluted by Fourier Transform (Muller et al) into a frequency signal from each ion. The frequencies are calibrated to known masses using reference molecules. Resolution from an Orbitrap is among the highest available at up to 100,000, as in the LTQ OrbitrapXL used in this research (and 240,000 in the most recent models)(ThermoFisherScientific, 2009b). The mass range is sufficient only for small molecules, peptides and small proteins at 50-5000 *m/z*. Orbitrap instruments are most frequently paired with an electrospray ionization source (ESI). ESI uses high voltage to vaporize

compounds in the liquid phase emitted from capillary, which is convenient for complex peptide mixtures and high-throughput analysis. In the work presented here, a high performance liquid chromatography (HPLC) in tandem with an Orbitrap Velos allows a complex peptide mixture to be separated based on hydrophobicity in sixty to ninety minutes per sample. A data-dependent MS/MS analysis is performed throughout the liquid chromatographic analysis. The "data" of the data-dependent analysis is a spectrum collected from the mix of parent compounds at high resolution in the Orbitrap. From these data, a single peak width of *m/z* is isolated and fragmented, and a spectrum of the mix of fragments is collected in the Orbitrap. The set of spectra produced from all peptides in a liquid chromatographic analysis can be submitted to the Mascot search engine to find matches in peptide masses and expected fragments. This software, the data-dependent MS/MS, the sample separation by liquid chromatography, and the high resolution and the simultaneous detection of peptides have become the standard for high throughput peptide and protein identification.

The triple quadrupole mass spectrometer is designed for selectivity and sensitivity (Hopfgartner et al, 2004). It is characterized by three sequential chambers that select a peptide, fragment the peptide, and select for a specific fragment. Each chamber consists of four parallel rods. Two opposing rods are positively charged and the other two rods are negatively charged. To keep the ions trapped, the rods switch polarity, inducing the ions to oscillate in the vacuum chamber. In addition to the polarity switching, a radio frequency voltage

(oscillating in intensity) can be applied to the rods. The oscillation frequency of a particular mass ion resonates with the radio frequency of the rods, causing that ion to increase in energy until it ejects itself from the chamber. With sequential ejections, a particular peak *m/z* can be isolated and retained in the quadrupole chamber. This isolation is performed up to two times: to isolate the parent and to isolate the fragment. A high number of ions in the instrument lowers sensitivity due to charge interference. Elimination of all molecules except at the *m/z* of interest increases sensitivity. In addition, the double selection allows for more confidence in the identity of the analyte. Triple quadrupole mass spectrometers are typically interfaced with HPLC sample separation. Data is frequently viewed as a chromatograph of the single parent/fragment isolation instead of a mass spectrum. An example of a triple quadrupole mass spectrometer is the Thermo TSQ Vantage (ThermoFisherScientific, 2009a). This instrument is used for high throughput quantitation of peptide abundance as well as abundance of small molecules such as drugs. Quantitation of a peptide is often performed using an identical synthetic peptide that is ^{13}C and ^{15}N isotope labeled at a lysine or arginine residue. These isotope labeled peptides will ionize with equal efficiency to the unlabeled analyte. The isotope label is spiked into the protein sample to standardize the sample preparation and the injection amount between the analyte and the isotope standard. Quantitation is then performed by comparing the ratio of the area under the curve of the analyte to the isotope labeled standard and the calibration curve of standards to the amount of standard spiked

into the protein sample. Alternatively, analyte peak area is compared to a calibration curve of a synthetic peptide that does not need to be labeled. While the triple quadrupole instrument is the standard for quantitative experiments, it is slow and inefficient at discovery of proteins in a complex sample. Similar to the software developed specifically for the Orbitrap, open source software "Skyline" has been developed for peptide analysis with triple quadrupole mass spectrometers, which facilitates peptide screening and quantitation of high-throughput samples (MacLean et al, 2010).

Some techniques are more readily accessible to biologists than mass spectrometry, including immunoblots and RT-PCR. Immunoblots can be used for protein or post-translational modification identification or quantitation. The technique can be very sensitive in a complex protein mixture, depending on the specificity of the antibody (Lee et al, 2000). Antibodies can be made to be very specific or promiscuous for multiple similar proteins (Michaud et al, 2003). In immunoblotting, the primary antibody binds the SDS-PAGE-resolved protein of interest and is detectable either directly, by conjugation to the enzyme horseradish peroxidase, or via conjugated secondary antibodies. The horseradish peroxidase on the antibody is then exposed to its substrate to create a fluorescent product only at the site of antibody binding. The choice of proteins to probe is limited by the antibodies that are available. Although the technique is costly in time and amount of sample if one is probing many proteins, the

technique is robust, inexpensive, and accessible to the traditional chemical biologist.

Studying mRNA expression is high-throughput and inexpensive, so it has become the surrogate for studying cellular status (Suter et al, 2004). The polymerase chain reaction (PCR) technique requires mRNA extracted from fresh tissue and a complementary primer sequence, so that one mRNA molecule is amplified to a detectable amount. Quantitation involves measuring the fluorescence intensity of the marker bound with the DNA products, compared to the fluorescence from tubes with known quantities of DNA. PCR is simple, although it does require some special instrumentation and careful technique. The quality of data depends on the quality of the original sample, which is very sensitive to contamination by an enzyme, RNase. Although mRNA abundance contains a substantial amount of information regarding the activity status of a cell, few RNAs perform the cellular activity itself and can sometimes be poorly representative of cellular activity (Vogel & Marcotte, 2012).

Cytochromes P450

The cytochromes P450 (CYP) are a family of enzymes that metabolize endogenous and exogenous compounds through oxygenation. Hydroxylated metabolites are more soluble and are more easily excreted. Of all metabolized drugs, CYPs are responsible for metabolizing about 75% (Wienkers & Heath, 2005). Each cytochrome P450 “isozyme” metabolizes multiple substrates. CYPs are located primarily on the endoplasmic reticulum, and they are unified by a

similar enzymatic mechanism for hydroxylation involving an iron-containing heme cofactor. Briefly, the donation of two electrons from two reductases and electron donors such as NADPH assist the CYP to activate the iron, then activate the oxygen molecule. This activation results in a highly unstable intermediate, which reacts with the substrate resulting in an oxidized product. The name “P450” comes from “pigment” (P) because the enzymes are colored red due to the heme, and from the wavelength of radiation emitted (450 nm) when carbon monoxide binds to the heme group instead of oxygen.

Activity of an enzyme can be determined in vitro by incubating enzyme with an energy source (NADPH) and substrate (Center for Drug Evaluation and Research (CDER)). The enzyme can be made recombinantly to investigate the activity of a single isozyme, or it can be isolated from tissue microsomes. Microsomes are protein-containing vesicles enriched for endoplasmic reticulum. They are prepared by homogenization of the tissue and ultracentrifugation of the organelle fraction. The resulting metabolites formed from the activity assay are analyzed for changes in their structure.

Nomenclature is based on homology using numbers and letters to indicate family, subfamily, and isozyme (e.g. CYP1A2). Family members exhibit greater than 40% sequence identity, subfamilies have greater than 59%, and isozymes represent gene clusters. Allelic variants have the same name if they are 97% identical. It is convention to capitalize the subfamily letter for human enzymes and use lowercase for mouse enzymes (Nebert et al, 1989) (Nelson & Strobel,

1987). For example, there exists a mouse Cyp1a2 and a human CYP1A2. Because they are within 97% homology the two isozymes have the same name except for the capitalization, which indicates the species of each allele.

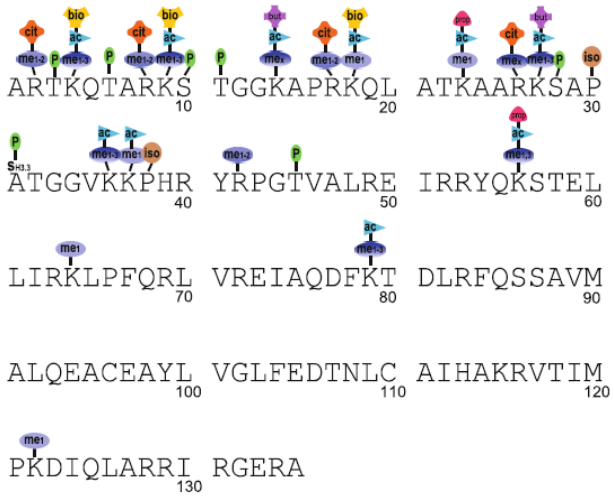
Despite mice serving as a drug metabolism model since 1940 (Lawrence et al, 1940), mouse Cyps being studied since the 1960's (Waterfield et al, 1969), and the genome being sequenced in 2002 (Waterston et al, 2002), the mouse Cyp proteome remains a mystery. There are 102 putatively functional murine Cyp genes (Nelson et al, 2004). Only recently has there been some proteomic characterization (using Orbitrap instruments) to determine which of these genes are expressed (Shinde et al, 2009; Sutton et al, 2010). Mouse Cyp research is otherwise limited to cross-reactivity of human protein-targeted antibodies, mRNA expression, and "knock-in" mouse models for human CYPs.

Histones

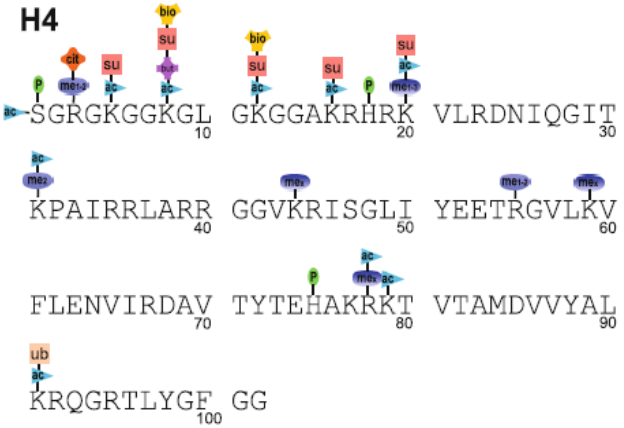
In the nucleus, the core histones H2A, H2B, H3 and H4 serve as a scaffold to wrap 146 basepairs of double-stranded DNA. The DNA-histone complex is called a nucleosome (Luger et al, 1997). Each histone has two general domains: an inner core region and an exposed N-terminal "tail", which can be post-translationally modified with acetylation, methylation, phosphorylation or other modifications (Cheung et al, 2000). This post-translational modification (PTM) profile of histones is dynamic, constantly remodified by acetyltransferase and deacetylase enzymes or methyltransferases and demethylases, which are in turn tightly regulated (Bonisch et al, 2008). Specific PTMs and combinations of PTMs

are proposed to function as a *histone code* that directs gene silencing, transcription, replication, cellular memory and gene recombination (Barlesi et al, 2007; Bonisch et al, 2008; Su et al, 2007). Elucidating histone post-translational modifications has thus become essential to understanding epigenetics. All known modifications on the four core histones are depicted in Figure 1-1.

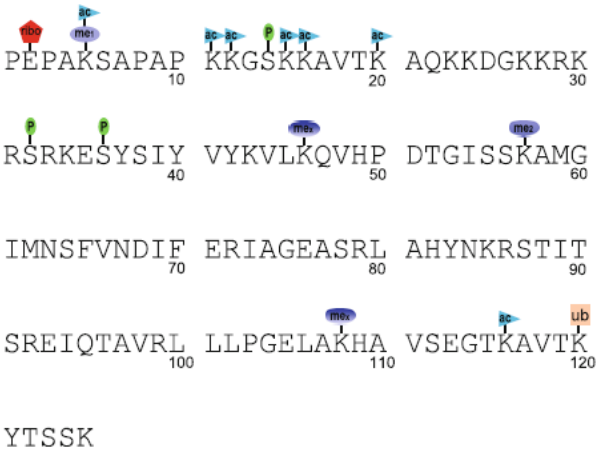
H3.1



H4



H2B



H2A

Figure 1-1. Reported human histone PTMs in the literature. (ac acetylation, me1 monomethylation, me2 dimethylation, me3 trimethylation, mex unspecified methylation number, P phosphorylation, ribo ADP ribosylation, ub ubiquitination, su sumoylation, bio biotinylation, iso proline isomerization, cit citrulination, but butylation). N-terminal methionine is cleaved in the functional protein, therefore numbering starts at 1 for the first non-methionine residue. Figure modified from (Young et al, 2010)

Histones are acetylated by histone acetyltransferases and deacetylated by histone deacetylases on lysine residues. Acetylation is correlated with transcriptional activation for example in younger populations. Deacetylation of histones is generally correlated with transcriptional silencing. Lysines can receive either an acetylation or 1-3 methyl groups and arginines can receive up to two methyl groups. Methylation typically acts as a repressor, but can also act as a transcriptional activator on certain histone residues. Histones are methylated by histone methyltransferases and demethylated by demethylases. Methyltransferases are hyperactive in cancer, for example. These and many other histone modifications represented in Figure 1-1 collectively contribute to chromatin regulation in the nucleus. Modification patterns of histones differ between species (Garcia et al, 2007a). Histones are heavily studied because they serve a fundamental role in biology, however they pose a technical challenge for mass spectrometrists.

A traditional “bottom up” proteomics method involves tryptic digestion of purified protein and analysis with mass spectrometry (Young et al, 2010). This technique assumes complete trypsinization and minimal lysine or arginine modifications for peptide quantitation. The “top-down” or “middle-down” peptide identification strategies used by Kelleher et al. (Siuti & Kelleher, 2007), along with lysine derivatization methods described by Garcia et al. (Garcia et al, 2007b; Garcia et al, 2007d), use intact or longer sequences by direct fragmentation via electron capture dissociation on a Fourier Transform Mass Spectrometer (Boyne et al,

2006; Jiang et al, 2007; Siuti et al, 2006). It is assumed that the number and/or position of basic residues for the intact proteins will have minimal effect, compared with smaller structures, on their chemistry and therefore on the opportunities for relative quantitation of the different isoforms (Pesavento et al, 2004; Young et al, 2009). Furthermore, all possible structural/positional combinations will be intact, hence hydrophilic interaction liquid chromatography can be combined with fourier transform mass spectrometer to identify and distinguish more than 150 differentially modified forms of histone H3.2 in HeLa cells using a top down approach (Siuti & Kelleher, 2007). This method is quantitative and the most biologically relevant, yet it is limited by the complexity of the sample and has been called impractical for histone PTM identification (Young et al, 2010).

An alternative approach uses propionic anhydride to form propionyl analogs that differ in mass from the acetylated peptides by 14 Daltons and obviate the need for isotopic corrections to the relative abundances (Garcia et al, 2007c). This approach has been used in a two step derivatization process, carried out before and after tryptic digestion, in which the second derivatization carries a stable isotope d_5 analog to enable quantitative comparisons between two samples (Plazas-Mayorca et al, 2009). The resulting peptides, however, produce dissimilar chromatographic behavior, ionization efficiency and quantitative accuracy.

The method presented in this thesis research is based on an approach first introduced by Catherine Smith (Smith et al, 2003) using deuteroacetylation.

Smith and colleagues used a mixture of deuterated (d_4) acetic acid and deuterated (d_6) acetic anhydride as the derivatization reagent, converting all endogenously unmodified lysines present in the histone sample to deuteroacetylated species, with a mass shift of +45 Da (Smith, 2005; Smith et al, 2003).

Deuteroacetylation is an in vitro chemical reaction performed by mixing dried protein with a 5:1 mixture of deuteroacetic acid and the chemically reactive deuteroacetic anhydride, pictured in Figure 1-2. The reaction occurs spontaneously over several hours: the nucleophilic histone lysine nitrogen attacks the carbonyl electrophile of a deuteroacetic anhydride molecule. The nitrogen is then covalently deuteroacetylated with an inert deuteroacetic acid byproduct.

The label-modified lysines are then chemically and chromatographically equivalent to their in vivo acetylated counterparts, but distinguishable by a three Dalton mass difference (+45 Da vs. +42 Da, respectively). All lysine residues are then equivalently “blocked” from cleavage by trypsin, cleavage occurs at arginine residues only, and the fragments generated are now larger in size, providing in some cases the ability to observe multiple modifications. From the mass spectra, three Dalton mass shifts are used to determine and quantitate endogenously acetylated (d_0) and in vitro deuteroacetylated (d_3) peptides. This technique is inexpensive, quantitative, high-throughput. An experiment detailing the PTM analysis of an individual histone sample requires about two hours of instrument

time, and a similar time for the subsequent data analysis. It provides high sequence coverage, site-specific PTM identifications, and accurate relative quantitation of acetylated peptides with high repeatability.

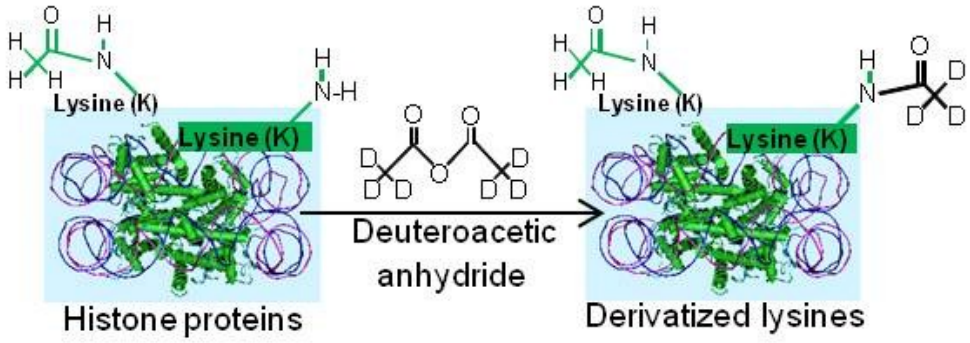


Figure 1-2. Deuteroacetylation tags all unmodified lysines with a deuteroacetyl group 3 Daltons heavier than endogenous acetylated lysines. Green bonds indicate endogenously formed bonds and black indicates in vitro modifications. Green indicates protein, purple indicates DNA.

In this thesis, I present the following contributions: 1) a proteomics SRM assay including 27 mouse cytochromes P450, 2) a semi-quantitative profile of mouse Cyp distribution in liver, kidney, lung, intestine, heart, brain, males of 3 ages, females of 3 ages, pregnant mice and fetuses, 3) characterization of efavirenz and its metabolites within the mouse model, including the effects of age, sex, and pregnancy status on metabolism, and 4) application of chemical derivatization for identification and quantification of post-translational modifications.

Chapter 2:

Murine Cytochrome P450 Quantification by Selected Reaction Monitoring

Introduction

Human CYP protein expression, mRNA expression, and CYP activity studies have revealed differences between tissue type, gender, age, and pregnancy status. For example, mRNA expression of a variety of human tissues demonstrates mainly hepatic expression of CYPs 1A2, 2C19, 2D6, and 3A4 (Bieche et al, 2007). Bieche et al. also identified predominant expression of certain Cyps in other tissues (e.g., CYP 2J2 in heart and 2F1 in lung). Age-dependent CYP expression has been described in immunoblotting studies, with perinatal expression of 3A7 and 2C9 and an increase in expression of 1A2, 2B6, 2D6, and 2E1 from 9 days old to 28 years old (Tateishi et al, 1997). One study employed an activity assay with human liver microsomes to identify a genetic basis for an apparent sex-dependent phenotypic discrepancy in CYP3A4 (Schirmer et al, 2007). Human heterogeneity in genotype and lifestyle also cause variability in CYP mRNA expression, resulting in 40- to 500-fold differences for some CYPs (Rodriguez-Antona et al, 2001). Researchers have used humanized mouse models to study human CYPs in an effort to reduce inter-individual heterogeneity. This approach enables the use of human CYP antibodies for protein detection, avoiding the need to raise antibodies to murine Cyps. For example, human 2E1 was expressed in mouse liver to evaluate the role of 2E1 induction in alcohol induced liver damage (Morgan et al, 2002).

Mice are the canonical academic model for drug efficacy and pharmacokinetic experiments yet only limited research has been done to characterize

endogenous Cyp protein expression in these model systems. A few prominent cases where murine Cyp proteins have been studied include Cyp2e1-null mice used to demonstrate a role for murine Cyp2e1 in high-fat mediated non-alcoholic steatohepatitis (Abdelmegeed et al, 2012) and Cyp4b1-null mice used to investigate its role in bioactivation of a pulmonary toxin (Parkinson et al, 2013). These experiments highlight the utility of murine Cyp mouse models for human disease and drug metabolism. Further characterization of murine Cyp protein expression can serve to promote the functional, mechanistic and biological relevance of Cyps.

To assess differences in protein levels in a more global high throughput fashion, we applied quantitative, targeted mass spectrometry to confidently detect a unique Cyp protein. SRM assays target dozens of peptides simultaneously with high specificity and sensitivity and can be applied to quantify proteins in any complex mixture.

The emerging field of “pharmacoproteomics” is expanding the understanding of molecular drug metabolism and Cyp biology, which has previously been driven by mRNA expression profiling and Cyp activity assays. Currently, less than a dozen publications apply mass spectrometry to Cyp protein biology. They report protein identification lists ranging from 11-37 Cyps (Flint et al, 2010; Jenkins et al, 2006) depending on sample preparation, mass spectrometer speed and sensitivity, and rigor of analysis. Patterson et al. have used global proteomics methods and optimized microsome preparation to improve the number of Cyp

isozyme identifications and quantitation accuracy (Sutton et al, 2010). Others have used global proteomics methods to identify changes in Cyp expression in mouse liver in response to psychological stress and exposure to carcinogens (Flint 2010) or to quantify Cyp induction by exposure to phenobarbitone (Jenkins 2006).

A mouse model characterizing natural Cyp biology would be valuable for comparison with current mouse models for pharmacokinetic and in vivo drug distribution experiments. The natural biological variation of Cyp expression in healthy mice is unexplored. To address this, we describe parallel Cyp proteomics and microsome activity assays. We identified the murine Cyps expressed in Balb/c mouse tissues and developed a 27-protein quantitative mass spectrometry assay. To characterize the influence of biological factors on Cyp expression, we applied the assay to liver, kidney, lung, intestine, heart and brain, from male and female mice at 3-4 weeks, 9-10 weeks, and 8-10 months of age. We found striking tissue specific expression differences, including the highest expression of cyp2b9 in the brain compared to other Cyps. Cyps 4b1, 2e1, and 2b9 showed a trend for higher expression in male kidney at all ages in comparison to female samples, whereas Cyp2d26 showed the opposite trend. The abundance of Cyps 4a10 and 4a12a correlated with age. Fetal liver was found to express Cyp 2d26 most abundantly in comparison to other Cyps, and pregnant mouse liver microsomes contained less Cyp 2e1 than other adult mice. Furthermore, we used the drug efavirenz to observe Cyp activity differences with

age, sex, and pregnancy status. These experiments allow direct comparison of proteomic results with activity of a sample xenobiotic substrate. This global assessment of murine Cyp expression provides a thorough and unique insight to murine Cyp biology.

Methods

Chemicals and Reagents. Efavirenz was purchased from Toronto Research Chemicals (Toronto, ON, Canada). NADPH was purchased from BD Biosciences (San Jose, CA). Iodoacetamide, phenylmethanesulfonyl fluoride, and urea were obtained from Sigma Aldrich (St. Louis, MO). Hexanes and Ethyl Acetate were purchased from Thermo Fisher Scientific (Sunnyvale, CA). All other chemicals were of the highest grade commercially available.

Mouse Tissue Preparation. Liver microsomes were prepared from the livers of male and female mice 3-4 weeks, 9-10 weeks, and 8-10 months of age in addition to pregnant mice (10-12 weeks of age) and fetuses (gestation day 16). Mice were euthanized as approved by the Johns Hopkins University Animal Care and Use Committee. The liver was excised and homogenized in a 1:4 (w/v) ratio of buffer (100 mM Tris HCl pH 7.4, 100 mM KCl, 1mM ethylenediaminetetraacetic acid, and 1mM phenylmethylsulfonylfluoride). The supernatant was filtered using a 70 μM nylon filter and tissue lysate was centrifuged at 10,000 x g for 30 minutes at 4°C. The supernatant was separated from cell debris then centrifuged at 10,000 x g for 30 minutes at 4°C.

Microsomes were isolated by ultracentrifugation at 100,000 x g, 4°C, for 90 minutes. The resulting pellet was washed using 3 ml 100mM Sodium Pyrophosphate buffer, pH 7.4 and centrifuged at 100,000xg 4°C for 60 min. Finally, the microsomes were resuspended in 200 µl of 50mM KPO4 buffer, pH 7.4 containing 0.1mM ethylenediaminetetraacetic acid, 0.1mM dithiothreitol, and 20% glycerol. Tissue total protein lysates were prepared from kidney, lung, intestine, heart and brain of the 10-12 weeks of age pregnant, 3-4 weeks, 9-10 weeks, and 8-10 months of age male and female mice. Fetal brain was the only other organ collected from fetus. All fetal brains from the pregnant mothers were pooled and lysed. Tissue was homogenized in lysis buffer containing 20 mM Tris HCl pH 7.5, 150 mM NaCl, 1mM ethylenediaminetetraacetic acid, 1 mM ethylene glycerol tetraacetic acid, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, 1mM Na_3VO_4, 1 ug/ml leupeptin, and 1mM phenylmethylsulfonylfluoride (Cell Signaling, Danvers, MA). Samples were centrifuged at 10,000 x g for 30 minutes at 4°C, and the supernatants were saved. The microsomes and lysates that were isolated from 4 individual mice within the same age and sex group were pooled. Fetal liver microsomes and brain lysates from all pregnant mice were pooled. Protein concentration was determined using the BCA Protein Assay Kit (Thermo Scientific Pierce, Rockford, IL), and protein was frozen at -80°C until use.

In-gel Peptide Isolation for Global Proteomics. Protein (25 µg) reduction and alkylation was performed in NuPAGE LDS sample buffer (Life Technologies,

Carlsbad, CA). NuPAGE Reducing Agent (Life Technologies, Carlsbad, CA) was added (50 mM DTT) and heated 70°C for 10 minutes. Samples were cooled to room temperature before incubation with of 1.2 µl of 1 M iodoacetamide (100 mM) for 20 minutes in the dark. NuPAGE Reducing Agent (80 mM) was added for continuation of the reaction for 20 minutes in the dark. A 10% tris-glycine gel (Biorad, Hercules, CA) was prepared with 25 µg protein loaded in each well. NuPAGE Antioxidant (500 µl, Life Technologies, Carlsbad, CA) was mixed in the running buffer. The gel was stained with SimplyBlue SafeStain (Life Technologies, Carlsbad, CA) according to manufacturer instructions. The 45-62 kDa band region of the gel was removed, corresponding to the range of molecular weights of murine Cyps. The stain was removed from the gel by incubation for 2 hours with 200 mM ammonium bicarbonate and 40% acetonitrile. Gels were dried under vacuum and trypsinized at 1:20 trypsin:protein ratio for 18 hours at 37°C. The resulting peptide solution was removed and the gel was extracted with 0.1% trifluoroacetic acid, 50 mM Ambic, and 50% acetonitrile by incubation at 37°C for 1 hour. Extracted peptides of each sample were combined with peptides for drying. Samples were reconstituted in 10µl of mobile phase A, 2 µl of which were injected for each LC-MS analysis.

Global Proteomics LC-MS Assay. Protein discovery experiments were performed by The Mass Spectrometry and Proteomics Facility at Johns Hopkins School of Medicine. Balb/c 9-10 weeks of age male and female tissue lysates (n=1) were analyzed to generate a list of proteins expressed in Balb/c mice. An

Eksigent nano 2-D HPLC was used in tandem with the high resolution Orbitrap Velos mass spectrometer (Thermo Scientific, Sunnyvale, CA). Peptides were separated on an in-house 10 cm column packed with 5 µm Magic C18 beads (Michrom Bioresource, Auburn, CA). Gradient solvents were composed of 2% acetonitrile 0.1% formic in LCMS grade water (Solvent A) and 90% acetonitrile 0.1% formic acid (Solvent B). Flow rate was set to 300 nL/min. The separation gradient was 90 minutes, followed by a 40 minute blank between each sample. The sample gradient increased from 2-8% to 5 minutes, 8%-32% to 70 minutes, 32-50% to 75 minutes, 50%-100% to 80 minutes, 100% to 84 minutes, and then to 2% to 90 minutes. Eluting peptides were sprayed into the mass spectrometer through a 10 µm emitter tip (New Objective, Woburn, MA). Peptides were analyzed at resolution 30,000 within 300–2000 *m/z*. The top 8 most abundant peptides were individually isolated at isolation width 1.9 Da, and fragmented (MS/MS) using collision energy 35. Fragments were scanned at resolution 7500 within 350-1800 *m/z*. Data-dependent acquisition was used for confident identification of as many peptides as possible. Dynamic exclusion of 30 seconds with a repeat count of 1 maximized the number of possible identifications. MIPS (monoisotopic ion precursor selection) was “on.” Lock mass was “on” (silocsane 371 Da). Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Proteome Discoverer (v1.3 Thermo Fisher Scientific). All MS/MS spectra were analyzed with Mascot (v.2.2 Matrix Science, London, UK) using the NCBI 167nr Database, *Musculus* species with acquired raw MS/MS data, trypsin

as the enzyme, 1 missed cleavage allowed, precursor mass tolerance 12 PPM, fragment mass tolerance 0.02 Da, and formation of y and b ions. Oxidation of methionine and carbamidomethyl of cysteine were specified in Mascot as variable modifications. For each sample, Mascot search result *.dat files with and without spectra extraction were processed in Scaffold to validate protein and peptide identifications.

In-solution Peptide Isolation for Targeted Proteomics. A proteomic SRM assay was developed to quantify the cytochrome Cyp proteins identified in the data-dependent mass spectrometry experiment. Samples for the targeted assay were prepared using the filter-aided sample preparation method developed by Matthias Mann (Wiśniewski, 2009). The targeted assay was applied to the previously described microsomes, including male and female Balb/c mice in 3 different age groups, plus pregnant mice and fetuses. In addition, pooled tissue lysates (N=4) of these mice were analyzed, including kidney, lung, intestine, heart and brain. Tissue lysates or microsomes (100 µg) were diluted in 50 µl of 50 mM ammonium bicarbonate. Proteins were reduced with NuPAGE Sample Reducing Agent (Life Technologies, Carlsbad, CA) for 60 minutes at 60°C. Free cysteines were alkylated in 100 mM iodoacetamide at room temperature for 15 minutes. Protein was washed with urea (9 M) three times with 300 µl on a 30 kDa filter (Sigma Aldrich, St. Louis, MO), followed by centrifugation at 14,000 x g for 10 minutes. Urea was removed using 5 washes of 25 mM ammonium bicarbonate (300 µl). Isotope labeled peptides were added (6000 fmol of 1a2

peptide and 600 fmol of 2b10, 2d22, 3a11, and 4v2 peptides from a 10:1:1:1:1 peptide ratio mix). Proteins were digested with 10 µg trypsin for 18 hours. An additional 1 µg fresh trypsin was added after 15 hours. The peptides were then collected by combining flow through and one 200 µl 25mM ammonium bicarbonate wash. Sample was dried under vacuum at 60°C. Each sample was resuspended in 40 µl mobile phase A directly before analysis. Each injection to the mass spectrometer contained 5 µg mouse protein (as calculated by the BCA assay performed on the microsome or lysate sample) and 30 fmol of 2b10, 2d22, 3a11, and 4v2 peptides and 300 fmol 1a2 peptide in a volume of 2 µl.

Targeted Proteomics LC-MS Assay. To write the mass spectrometric method and view the targeted method data, the open source program Skyline (MacCoss Lab, University of Washington School of Medicine) was employed. Cyp protein sequences obtained from the National Center for Biotechnology Information Reference Sequence Database (NCBI RefSeq Database) were imported into Skyline in the FASTA format. Skyline performed an in silico digestion of the proteins at lysines and arginines, allowing 0 missed cleavages, +2 and +3 charges, and length of 5 to 25 residues. Basic Local Alignment Search Tool (NCBI BLAST) was used to determine uniqueness of a peptide to the protein. Multiple rounds of transition selection were performed, prioritizing peptides that do not contain ragged ends, methionines, cysteines, glutamines and asparagines. One peptide for quantification and one peptide for protein confirmation were chosen for the final assay. The four transitions with the

highest signal to noise ratio were retained for the peptide intended for quantification; replicate analysis of the sample allowed for confident retention time identification. Five representative peptides were chosen for absolute quantitation. All peptides being monitored had either an isotope labeled or unlabeled peptide synthesized as a standard. Heavy labeled peptides (C-terminal (^{13}C)/(^{15}N) were synthesized by New England Peptide (Gardner, MA), and unlabeled peptides were synthesized by Sigma Biosciences (Rockville, MD).

Mass spectrometry and HPLC parameters for the quantitative proteomics assay were determined separately from the peptide and transition selection in Skyline (MacLean et al, 2010). Liquid chromatography was performed by a nanoAcquity ultra-performance liquid chromatograph (Waters, Wexford, Ireland) in tandem with a triple quadrupole TSQ Vantage mass spectrometer (Thermo Scientific, Sunnyvale, CA). Peptides were separated using a Halo reverse phase C18 column with 2.7 µm bead diameter and dimensions 2.1 x 100 mm (Mac-Mod Analytical, Chadds Ford, PA). Column was heated to 40°C, and samples were kept at 6°C. Flow rate was 100 µl/min. Mobile phase A was 0.1% formic acid in water and mobile phase B as 0.1% formic acid in 90% acetonitrile. The 30 minute chromatography sequence started at 5% B for 2 minutes, increased to 60%B to 26 minutes, 98% to 27 minutes, 5% to 28 minutes, and remained at 5% to 30 minutes. Cycle time was 5 seconds, peak widths for Q1 and Q3 were 0.70 FWHM, and Chromatographic Filter Peak Width was set to 50 seconds. All ions

were monitored in positive mode and analyzed by software Thermo XCalibur (version 2.1; Sunnyvale, CA).

Proteomics Quantitation. The isotope labeled peptide standards were prepared in triplicate by spiking commercial Balb/c mouse liver microsomes, following the normal sample preparation described above, with dilutions of synthetic isotope labeled peptide mix. Isotope labeled peptide mix was serially diluted from 20 nM to 200 μM resulting in a final calibration curve spanning 5 orders of magnitude (1.5 fmol, 3 fmol, 15 fmol, 30 fmol, 75 fmol, 150 fmol, 300 fmol, 1.5 pmol, 3 pmol, 15 pmol). Quantification was performed using the peptide transition with highest signal to noise ratio in the extracted chromatograph of each peptide. Calibrant peak areas were normalized to the 30 fmol peak area and fit to a $1/y^2$ weighted linear regression in GraphPad Prism. Analyte concentration was calculated from the analyte/isotope labeled peptide area under the curve ratio using the linear regression equation from the normalized isotope calibrants. The calibration curves for the synthetic unlabeled peptides were prepared by serial dilution of the equimolar peptide mix from a 500 μM stock in triplicate. Direct injection (2 μl) of the dilutions produced calibration points spanning 6 orders of magnitude (1.5 fmol, 3 fmol, 15 fmol, 30 fmol, 75 fmol, 150 fmol, 300 fmol, 1.5 pmol, and 3 pmol). At least 6 of the above dilutions were used for each peptide calibration curve according to linear range except the peptides representing proteins 2c37 and 4a10, which contained 5 points. Unlabeled peptide standards were averaged (n=3) and fit to a $1/y^2$ weighted linear regression in GraphPad Prism (San Diego,

CA). Relative quantification of peptides was performed using the analyte peptide peak area and the regression equation to calculate the approximate concentration of peptide in the tissue.

Microsome Assay. Differences in Cyp activity of the in-house prepared pooled microsomes were measured by the microsome activity assay. Total volume per reaction was 250 µl, containing 1 mg/ml pooled liver microsomes (20 mg/ml), efavirenz (7 µM), potassium phosphate buffer pH 7.4 (0.1 M), and water. The reaction was initiated at 37°C in a water bath by the addition of NADPH-regenerating system (BD Biosciences, San Jose, CA) and incubated 30 minutes at 37°C. The reaction was quenched with the addition of 250 µl 1:1 hexanes: ethyl acetate. The samples were then placed on ice 10 minutes followed by centrifugation at 10,000 x g and 4°C. The organic layer was saved, and the aqueous protein fraction underwent a second extraction with 250 µl hexanes: ethyl acetate. The organic fractions were combined and dried under vacuum at 60°C followed by resuspension in 20 µl methanol and 10 µl injection for LC-MS analysis. A fluorinated efavirenz analog was used as an internal standard for quantification. A racemic 6-fluorinated analog of efavirenz was synthesized by Dr. David Meyers (Avery et al, 2010; Avery et al, 2013) using modifications of previously published meth. F-EFV was added after the reaction at a final concentration of 1 ng/µl. Signal intensities of 8-hydroxyefavirenz and 5-hydroxyefavirenz metabolites were normalized by the intensity of the F-EFV peak intensity.

Efavirenz Metabolite LC-MS Assay. Determination of metabolite formation was performed using a Dionex Ultimate 3000 ultra high-performance liquid chromatograph (Thermo Scientific, Sunnyvale, CA) in tandem with a triple quadrupole TSQ Vantage mass spectrometer (Thermo Scientific, Sunnyvale, CA). Samples were injected onto an XTerra reverse phase column, with 2.5µm C18 beads and 2.1 x 50 mm dimensions (Waters, Milford, MA). Column and samples remained at room temperature during analysis. Sample was injected by the Dionex autosampler at a draw speed of 5 µl/s, a dispense speed of 20 µl/s, followed by inject wash of 100 µl and loop wash by a factor of 2. Flow rate was 0.4 ml/min, with mobile phase A as 0.1% formic acid in water and mobile phase B as 0.1% formic acid in acetonitrile. The column was conditioned at 35% B. From 0.2 minutes to 5 minutes, mobile phase B increased linearly to 60%. Remaining sample was eluted at 95% B for 0.8 minutes, and the column equilibrated at 5% B for 0.2 minutes. Efavirenz and fluorinated efavirenz were monitored using MRMs at *m/z* of precursor/product ion pairs 313.85/243.98 and 313.85/243.98, respectively. Scan width was 0.002, scan time was 0.001, peak widths for Q1 and Q3 were 0.70 FWHM, and collision energies were 18 and 20, respectively. Mono- and Di- hydroxylated efavirenz metabolites were monitored in product scanning mode. The expected *m/z* of the parent ions were 329.821, 345.789, and 361.765, respectively. Scan time was 0.1, peak widths for Q1 and Q3 were 0.70 FWHM, and collision energies were 14, 17, and 14, respectively.

Product ions were scanned from *m/z* 50 to 370. All ions were monitored in negative ionization mode.

Statistical Analysis. Graphs and the two-tailed unpaired t tests were performed using GraphPad Prism version 6.02 for Windows (GraphPad Software Inc., San Diego, CA).

Results

Two mass spectrometric methods were employed for identification (global approach) and quantification (targeted approach) of murine cytochrome P450s (Figure 2-1). Firstly, for unbiased detection of mouse Cyp expression, we performed a global proteomics experiment on liver, kidney, lung, intestine, heart, and brain lysates from 9-10 weeks of age male and female mice. Lysates were desalted by gel electrophoresis, and the mass region containing Cyps was selected for increased coverage of Cyps specifically. Proteins were digested with trypsin in the gel, peptides were separated by nano HPLC and analyzed by high resolution Orbitrap MS/MS. Mascot analysis revealed 22 cytochrome P450 enzymes in at least one organ of one mouse with 95% confidence containing at least 2 peptides.

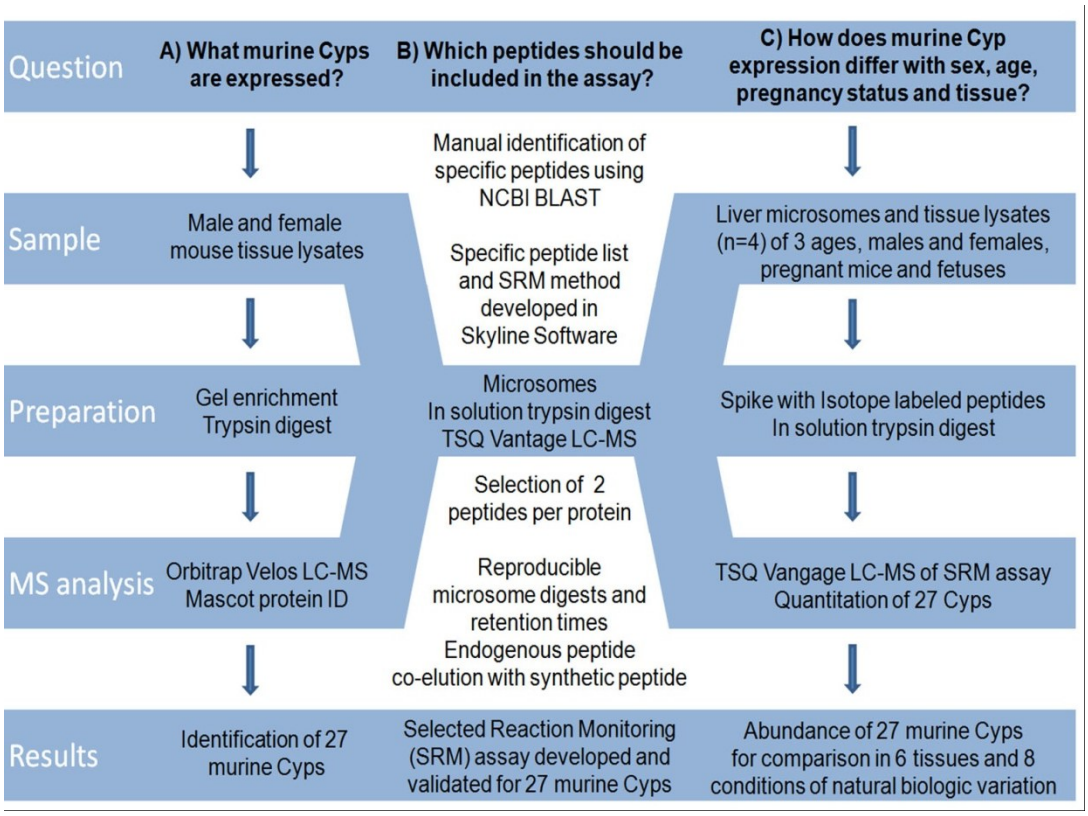


Figure 2-1. Mass Spectrometry workflows for identification and quantitation of murine Cyps. A proteomics approach was applied to murine Cyp biology. The discovery approach was used to A) compile a list of Cyps expressed in mouse liver, kidney, lung, heart, intestine, and brain using a proteomics discovery methods, B) develop a SRM assay for the proteins identified, and C) use the SRM assay developed to apply a targeted approach to quantitate Cyp expression in mouse liver microsomes and tissue lysates.

The proteins identified in female (Figure 2-2A) and male (Figure 2-2B) lysates are shown, with the number of confident spectra, “spectral counts,” binned and color coded according to detection frequency. Results demonstrated that liver readily expresses several known prominent murine Cyps, including Cyps 1a2, 2b10, 3a11, 2c39 and 2d22 and several Cyps not previously associated with liver expression Cyps 2u1, 4a10, 4a12a, 4b1, and 4v2. The spectral count of proteins represents detection of all peptides possible from each protein. Cyps 2b9, 2d9, 2j5, 3a13, and 4v2 were identified in a single tissue of either male or female lysate, reflecting the importance of using both sexes and multiple tissues for murine Cyp identification. Cyps expressed in intestine, heart, and brain were undetectable with the protein discovery method, although global proteomics experiments are known to be less sensitive than targeted assays. Five more murine Cyp proteins were included in the list for targeted assay development because they were identified in mouse microsomes of previous experiments (data not shown). Global proteomics experiments resulted in a total list of 27 murine proteins.

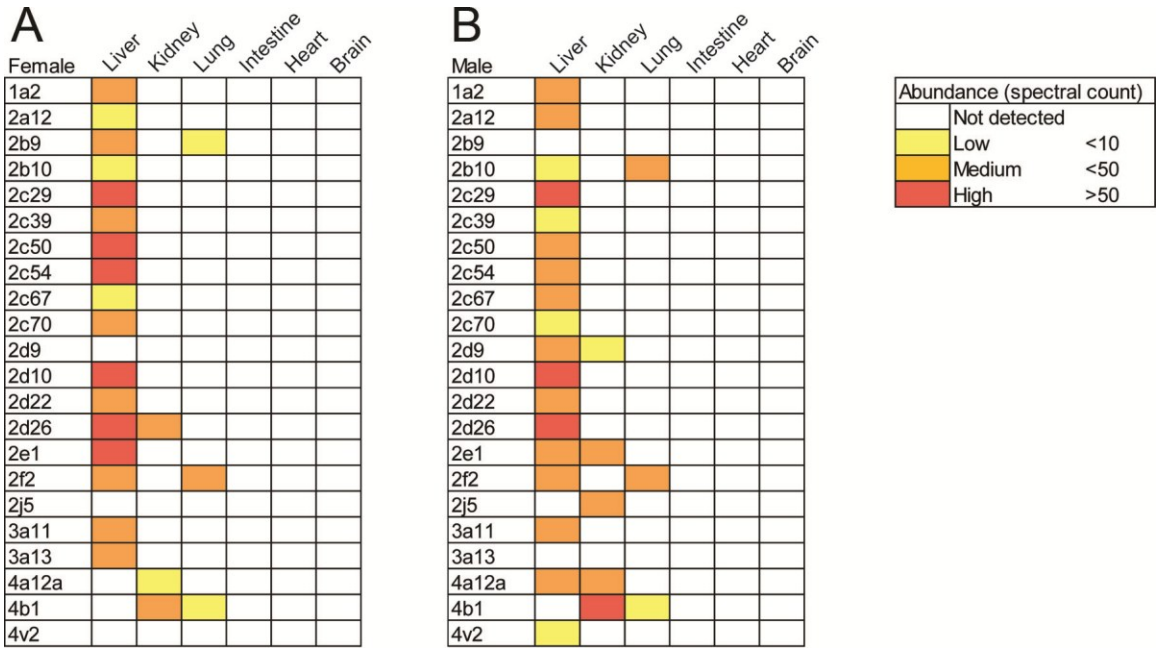


Figure 2-2. Cytochrome P450 expression. Liver, kidney, lung, intestine, heart, and brain of one (A) female and one (B) male mouse of 9-10 weeks age were lysed in the presence of protease inhibitors. Cyp enzymes were reduced and alkylated, then washed using gel electrophoresis. The mass region corresponding to the range of Cyp masses, between 45 kDa and 62 kDa, was cut and trypsinized. Peptides were submitted to the Mass Spectrometry Core Facility at Johns Hopkins School of Medicine for analysis by Orbitrap Velos high resolution mass spectrometer with confidence set at a minimum of 2 peptides. Results were searched by mascot and visualized in the software Scaffold. Cytochromes P450 that were not identified in these samples are listed because they were identified in female 9 weeks of age, fetus, or on the TSQ Vantage mass spectrometer. All spectra represented in the spectral count are of 95%.

The peptide and transition chosen to represent each isozyme for quantitation are shown in Table 2-1 and all transitions shown in Table 2-2. The full assay was validated for repeatability in sample preparation, chromatographic retention time, and transition linearity for quantitation. Raw data is visualized as concentric chromatograph traces of the peptide transitions, as shown in Figure 2-3. Area under the curve of the most abundant transition is used for quantification. Absolute quantification using the targeted assay resulted in limits of quantification of 30, 0.6, 3, 3, and 3 fmol isotope labeled peptide per µg microsome for Cyps 1a2, 2b10, 2d22, 3a11, and 4v2, respectively. Reported abundances are calculated based on 5 µg of lysate injected per assay. The limits of quantification of the remaining 22 peptides had a similar range as the isotope labeled peptides.

Table 2-1. SRM assay design for tryptic peptides and isotope labeled peptides. The cytochrome P450s, protein accession numbers, peptide sequences and transitions used for quantitation in the targeted mass spectrometric assay are outlined below. The protein list was generated from a global proteomics analysis of protein identified in Balb/c 9-10 weeks of age male and female tissues. All peptides selected are specific to the protein except for the Cyp 4a10 peptide, which is identical in Cyp4a32.

Iso-zyme	Accession no.	Peptide sequence	MRM (parent/transition)
1a2	NP_034123.1	T_{254}FNDNFVLFLQK_{265}	743.390/747.476 (y6)
2a12	NP_598418.1	E_{88}ALVDHAEEFSGR_{100}	730.344/1047.449 (y9)
2b9	NP_034130.1	D_{263}YIDTYLLR_{271}	586.303/780.425 (y6)
2b10	NP_034129.1	G_{99}TVAVVEPTFK_{109}	574.322/819.461 (y7)
2c29	NP_031841.3	N_{49}ISQSFTNFSK_{59}	636.815/830.404 (y7)
2c37	NP_034131.2	Y_{308}AILLLLK_{315}	473.820/712.533 (y6)
2c39	NP_034133.2	F_{126}TLTTLR_{132}	426.253/603.382 (y5)
2c50	NP_598905.2	G_{384}TNVITSLSSVLR_{396}	673.886/862.499 (y8)
2c54	NP_996260.1	E_{85}ALVDHGDVFAGR_{97}	693.344/973.449 (y9)
2c67	NP_001019890.1	V_{125}FTINTLR_{132}	482.285/503.294 (y4)
2c70	NP_663474.2	E_{85}ALIDQGDEFSDK_{97}	733.836/1040.417 (y9)
2d9	NP_034136.2	D_{408}ESVWEKPLR_{417}	629.825/828.473 (y6)
2d10	NP_034135.2	F_{369}GDIAPLNLPR_{379}	606.840/709.436 (y6)
2d22	NP_062797.3	T_{263}TWDPTQPPR_{272}	599.796/695.383 (y6)
2d26	NP_083838.1	G_{121}VILAPYGPEWR_{132}	679.367/454.302 (b5)
2e1	NP_067257.1	G_{113}IIFNNGPTWK_{123}	623.833/963.468 (y8)
2f2	NP_031843.2	S_{48}QDLLTSLTK_{57}	553.309/662.408 (y6)
2j5	NP_034137.1	L_{50}PFVGNFFQIDTK_{62}	763.406/1012.510 (y8)
2u1	NP_082092.2	E_{464}TFIPFGIGK_{473}	554.805/618.361 (y6)
3a11	NP_031844.1	L_{332}QDEIDEALPNK_{343}	692.851/358.208 (y3)

3a13	NP_031845.1	D_{244}VISFFTTSVER$_{255}$	700.856/839.426 (y7)
3a25	NP_062766.2	F_{447}ALISIK$_{453}$	396.255/573.397 (y5)
4a10	NP_034141.3	T_{215}YLQAIGDLNNLFHSR$_{230}$	621.323/887.448 (y7)
4a12a	NP_803125.2	S_{214}YIQAVEDLNDLVFSR$_{229}$	934.973/1207.595 (y10)
4b1	NP_031849.1	G_{125}LLVLEGPK$_{133}$	463.289/642.382 (y6)
4f13	NP_570952.1	S_{453}PLAFIPFSAGTR$_{465}$	682.372/735.378 (y7)
4v2	NP_598730.1	V_{391}FPSVPLFAR$_{400}$	566.829/886.515 (y8)

Table 2-2. SRM assay transition list. The cytochrome P450s and transitions monitored for the targeted mass spectrometric assay are outlined below. The three highest intensity transitions per peptide monitored were included in the final assay. The three transitions were used for qualitative assessment of the limit of detection, based on signal to noise ratio of the three transitions and consistent presence in replicates. MRMs are represented as parent/fragment in the order of highest transition *m/z* to lowest. Proteins with a star are heavy labeled (^{13}C and ^{15}N) synthetic peptides included in the assay for quantitation.

Iso-zyme	MRM1 (transition)	MRM2 (transition)	MRM3 (transition)
1a2	743.390/1008.588 (y8)	743.390/894.545 (y7)	743.390/747.476 (y6)
1a2*	747.398/1016.602 (y8)	747.398/902.559 (y7)	747.398/755.491 (y6)
2a12	730.344/1047.449 (y9)	730.344/932.422 (y8)	730.344/795.363 (y7)
2b9	586.303/1056.572 (y8)	586.303/893.509 (y7)	586.303/780.425 (y6)
2b10	574.322/819.461 (y7)	574.322/621.324 (y5)	574.322/500.296 (y4)
2b10*	578.329/827.475 (y7)	578.329/629.338 (y5)	578.329/500.296 (y4)
2c29	636.815/1045.495 (y9)	636.815/830.404 (y7)	636.815/743.372 (y6)
2c37	473.820/783.57 (y7)	473.820/712.533 (y6)	473.820/599.449 (y5)
2c39	426.253/704.430 (y6)	426.253/603.382 (y5)	426.253/490.298 (y4)
2c50	673.886/975.583 (y9)	673.886/862.499 (y8)	673.886/761.452 (y8)
2c54	693.344/973.449 (y9)	693.344/858.422 (y8)	693.344/721.363 (y7)
2c67	482.285/616.378 (y5)	482.285/503.294 (y4)	482.285/247.144 (b2)
2c70	733.836/1153.501 (y10)	733.836/1040.417 (y9)	733.836/797.331 (y7)
2d9	629.825/828.473 (y6)	629.825/642.393 (y5)	629.825/385.256 (y3)
2d10	606.840/780.473 (y7)	606.840/709.436 (y6)	606.840/504.245 (b5)
2d22	599.796/695.383 (y6)	599.796/810.410 (y7)	599.796/369.224 (y3)
2d22*	604.800/705.392 (y6)	604.800/820.419 (y7)	604.800/379.233 (y3)
2d26	679.367/975.468 (y8)	679.367/904.431 (y7)	679.367/454.302 (b5)

2e1	623.833/963.468 (y8)	623.833/816.400 (y7)	623.833/702.357 (y6)
2f2	553.309/890.519 (y8)	553.309/775.492 (y7)	553.309/662.408 (y6)
2j5	763.406/1069.531 (y9)	763.406/1012.510 (y8)	763.406/898.467 (y7)
2u1	554.805/878.513 (y8)	554.805/731.445 (y7)	554.805/618.361 (y6)
3a11	692.851/1143.553 (y10)	692.851/786.399 (y7)	692.851/358.208 (y3)
3a11*	696.858/1151.567 (y10)	696.858/794.413 (y7)	696.858/366.223 (y3)
3a13	700.856/1073.526 (y9)	700.856/986.494 (y8)	700.856/839.426 (y7)
3a25	396.255/644.434 (y6)	396.255/573.397 (y5)	396.255/460.313 (y4)
4a10	621.323/862.431 (b8)	621.323/1000.532 (y8)	621.323/887.448 (y7)
4a12a	934.973/1207.595 (y10)	934.973/1078.553 (y9)	934.973/963.526 (y8)
4b1	463.289/755.466 (y7)	463.289/642.382 (y6)	463.289/543.314 (y5)
4f13	682.372/995.531 (y9)	682.372/848.462 (y8)	682.372/735.378 (y7)
4v2	566.829/886.515(y8)	566.829/702.430 (y6)	566.829/603.361 (y5)
4v2*	571.833/896.523 (y8)	571.833/712.438 (y6)	571.833/613.370 (y5)

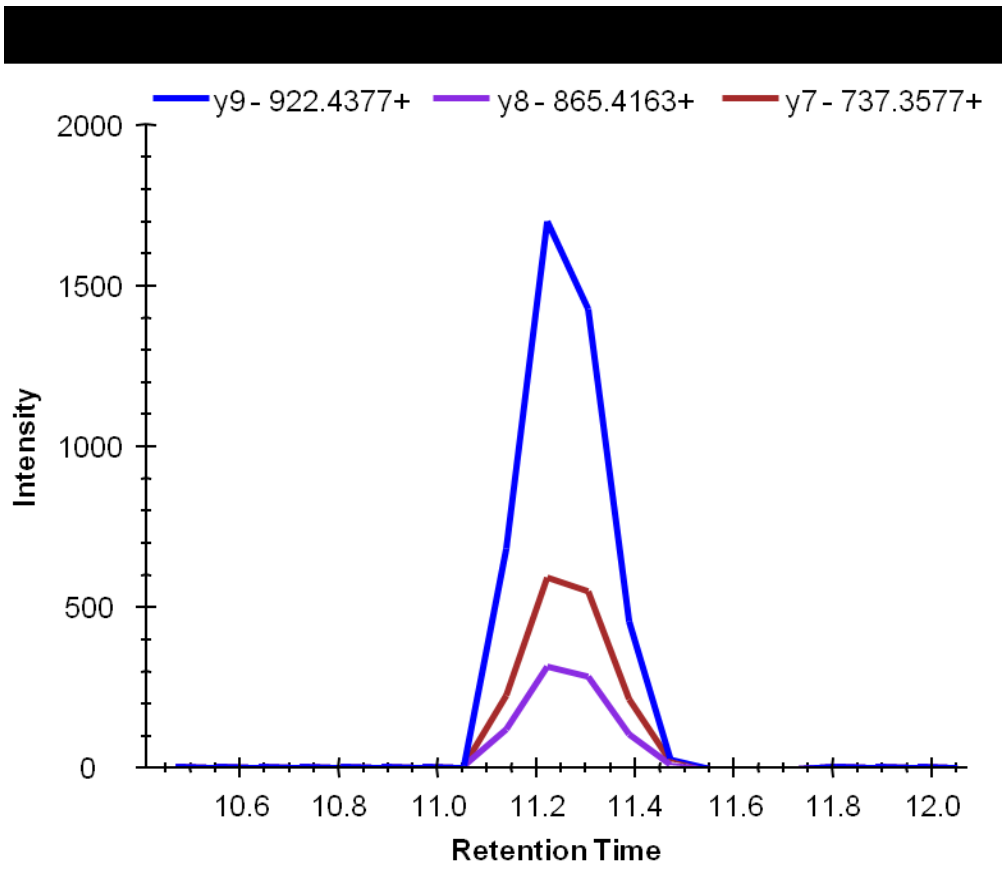
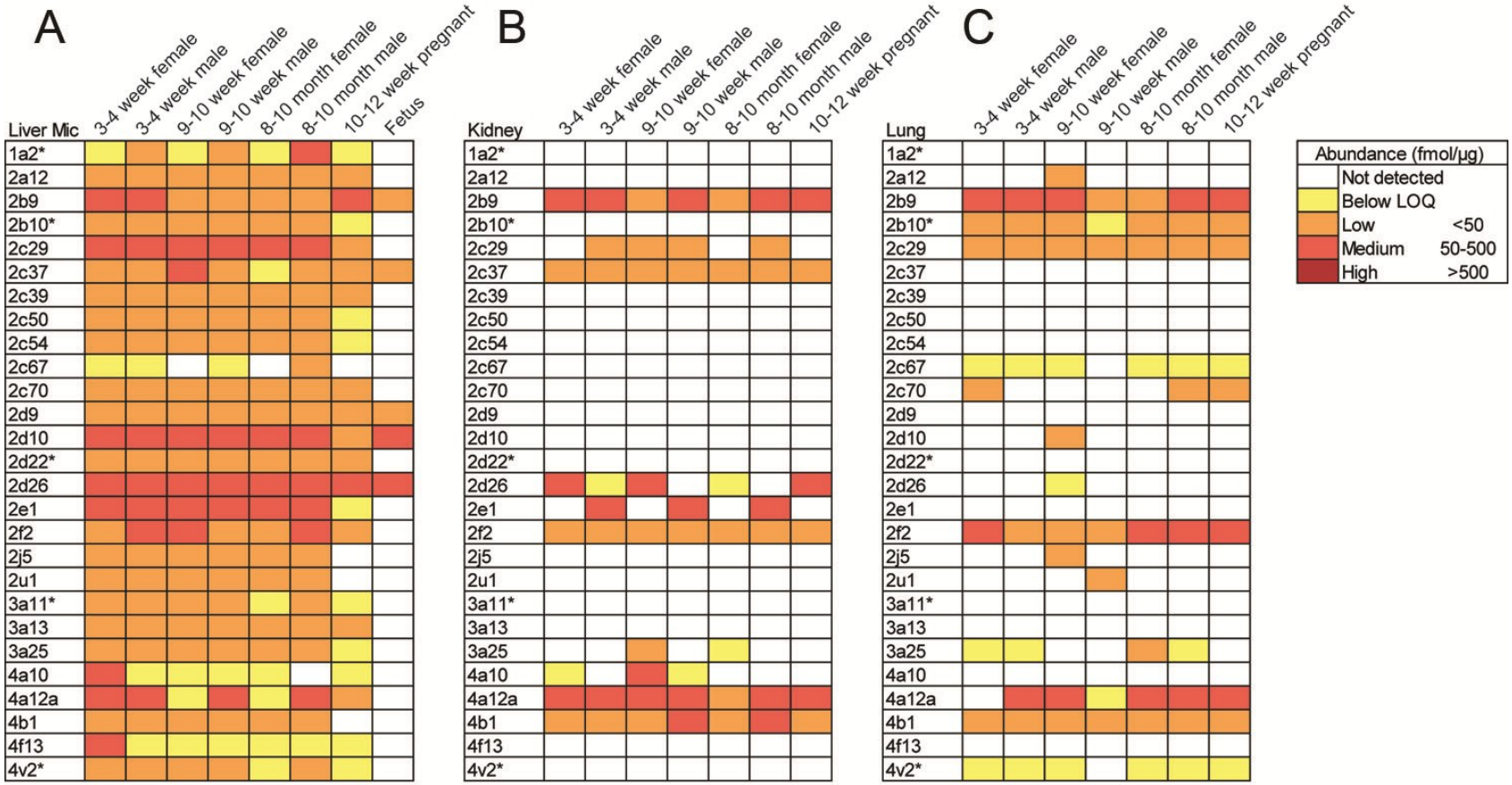


Figure 2-3. SRM chromatograph. The chromatograph represents the 2b10 peptide ATLDPSVPR eluting from a C18 column between 11.0 and 11.5 minutes. After the peptide *m/z* is isolated, it undergoes fragmentation, and a single fragment is isolated and detected. The *m/z* of the fragment ions are indicated by colored traces in the chromagraph. Intensity represents the relative abundance of the peptide.

The resulting targeted quantitative mass spectrometry method was applied to liver microsomes from 3-4 week, 9-10 week, and 8-10 months of age male and female mice (Figure 2-4A). In addition, microsomes from 10-12 week old pregnant mice and mixed sex fetuses were screened (Figure 2-4A). Brain lysates from all mice were analyzed (Figure 2-4D). Kidney (Figure 2-4B), lung (Figure 2-4C), intestine (Figure 2-4E), and heart (Figure 2-4F) lysates from the male and female 3-4 week, 9-10 week, and 8-10 month and 10-12 weeks of age pregnant mice were also analyzed for Cyp expression.


A
B
C
Liver Mic
Kidney
Lung
3-4 week female
3-4 week male
9-10 week female
9-10 week male
8-10 month female
8-10 month male
10-12 week pregnant
Fetus
1a2*
2a12
2b9
2b10*
2c29
2c37
2c39
2c50
2c54
2c67
2c70
2d9
2d10
2d22*
2d26
2e1
2f2
2j5
2u1
3a11*
3a13
3a25
4a10
4a12a
4b1
4f13
4v2*
Abundance (fmol/µg)
Not detected
Below LOQ
Low <50
Medium 50-500
High >500

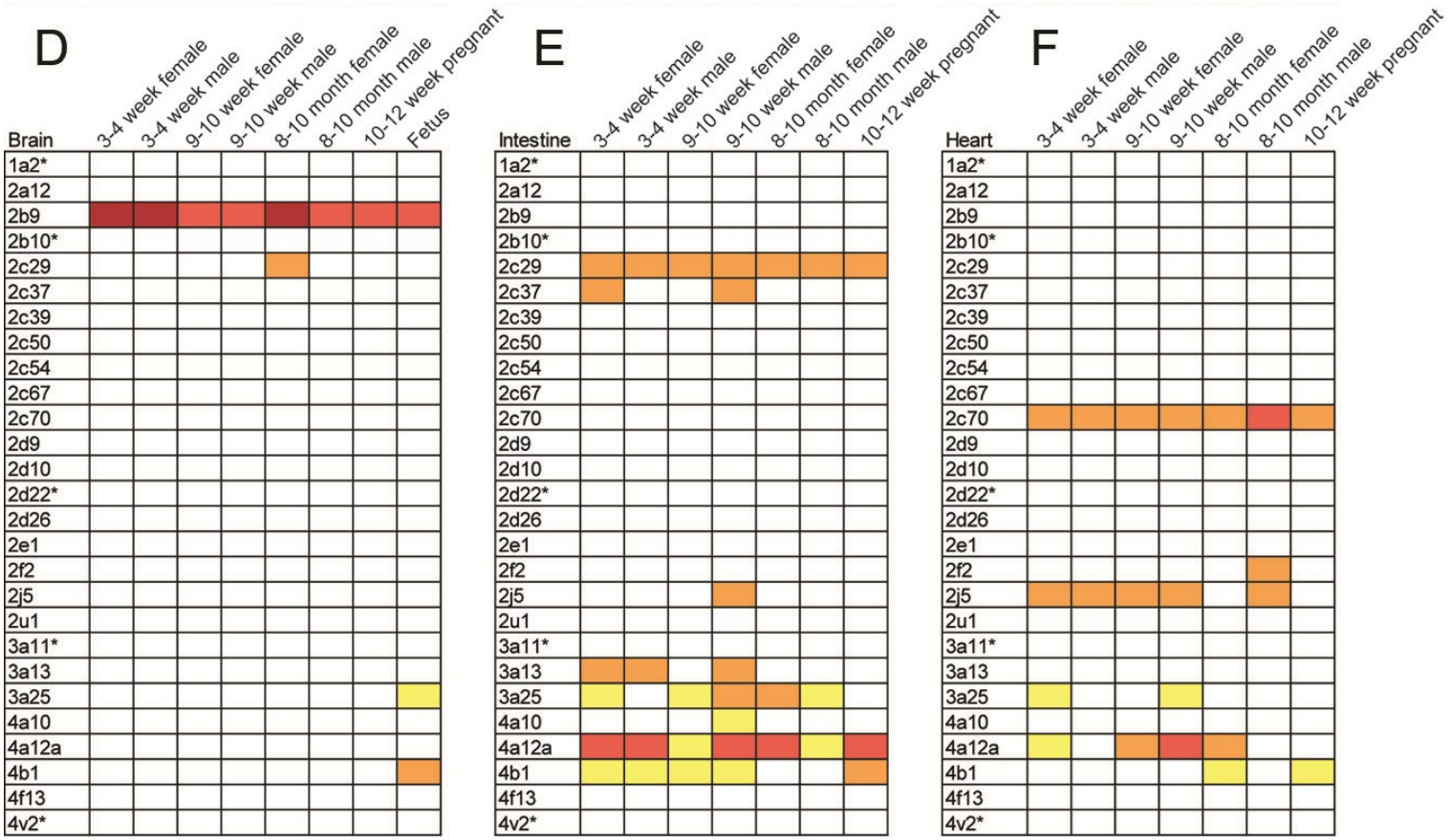

D
Brain
3-4 week female
3-4 week male
9-10 week female
9-10 week male
8-10 month female
8-10 month male
10-12 week pregnant
Fetus
1a2*
2a12
2b9
2b10*
2c29
2c37
2c39
2c50
2c54
2c67
2c70
2d9
2d10
2d22*
2d26
2e1
2f2
2j5
2u1
3a11*
3a13
3a25
4a10
4a12a
4b1
4f13
4v2*
E
Intestine
3-4 week female
3-4 week male
9-10 week female
9-10 week male
8-10 month female
8-10 month male
10-12 week pregnant
1a2*
2a12
2b9
2b10*
2c29
2c37
2c39
2c50
2c54
2c67
2c70
2d9
2d10
2d22*
2d26
2e1
2f2
2j5
2u1
3a11*
3a13
3a25
4a10
4a12a
4b1
4f13
4v2*
F
Heart
3-4 week female
3-4 week male
9-10 week female
9-10 week male
8-10 month female
8-10 month male
10-12 week pregnant
1a2*
2a12
2b9
2b10*
2c29
2c37
2c39
2c50
2c54
2c67
2c70
2d9
2d10
2d22*
2d26
2e1
2f2
2j5
2u1
3a11*
3a13
3a25
4a10
4a12a
4b1
4f13
4v2*

Figure 2-4. Cytochrome P450 quantitation in liver microsomes and tissue lysates comparing age, sex, and pregnancy status. Liver microsomes (A) and (B) kidney, (C) lung, (D) brain, (E) intestine, and (F) heart were prepared from Balb/c pregnant mice, fetuses, and male and female 3-4 week, 9-10 week, and 8-10 months of age mice. Average area under the curve for the three replicates is indicated by color. Cyp enzymes were reduced and alkylated, washed with 9M urea on a 30 kDa filter, and trypsinized in a 25mM ammonium bicarbonate buffer. Proteins were analyzed by the targeted SRM assay using the TSQ triple quadrupole mass spectrometer. Absolute quantitation was performed on 1a2, 2d22, 3a11, 2b10, and 4v2, indicated by a star in the Isozyme column. Quantification was performed in Skyline, and peak signal to noise ratio was manually validated.

All 27 proteins included in the targeted screen were identified in liver microsome samples (Figure 2-4A). Cyps 4b1, 3a25, 4a12a, and 2c29 were expressed most frequently across all organs. The most abundant proteins were Cyps 2c29, 2d26, and 2d10. Cyp 1a2 and 4a12a appeared to be less abundant in microsomes from female mice, and Cyp2e1 appeared to be less abundant in pregnant mice. For most Cyps abundance in the pregnant liver was comparable or slightly less than the age matched female mouse, with the exception of Cyp2b9, with higher expression in the pregnant mouse. Cyp2b9 also showed higher expression in 3-4 weeks of age mice in comparison to mice of other ages. Cyps 4a10 and 4f13 were detected in high abundance in 3-4 weeks of age female mouse microsomes, while abundance was below the limit of quantitation at all other ages. Fetal liver expressed Cyps 2b9, 2c37, 2d9, 2d10, and 2d26.

Cyp 2b9 was detected in the highest abundance in brain tissue (Figure 2-4D). Similar to trends in liver microsomes, cyp2b9 was found in higher abundance in the 3-4 weeks of age male and female mouse brain lysates compared to adult levels. Pregnant mice Cyp2b9 brain expression was comparable to the other adults. Cyps 2b9, 3a25, and 4a12a were detected in fetal brain lysate. Cyp2c29 was confidently detected in the 8-10 months female brain lysate.

The targeted proteomics assay identified 10 of the 27 proteins in kidney lysates (Figure 2-4B). Cyps 4a12a, 2b9, and 4a10 demonstrated the highest expression. Kidney was the only extrahepatic organ expressing detectable levels of 2e1. Male-specific 2e1, 4b1 and 2b9 expression in kidney was identified at all three ages. Cyp2e1 was detected in high abundance in all male kidney lysates, but

was not detected in any of the 4 female kidney lysates. Male Cyp4b1 expression was consistently ten-fold higher than female expression at all ages including pregnant mice. Cyp2b9 was detected in male kidney lysates of all ages but only one of the four female kidney lysates. Conversely, Cyp3a25 was only detected in adult female kidney lysates. The largest number of Cyps exhibiting sex-specific trends was found in the kidney.

The number of Cyps identified in lung tissue was second only to liver expression: 15 of the 27 proteins screened were detected in lung tissue (Figure 2-4C). Cyps 4a12a, 2f2, 2b9, and 2b10 were amongst the most highly expressed Cyps. In contrast to the kidney, no sex-, age-, or pregnancy-specific trends were identified in lung tissue. In comparison to the other extrahepatic organs it appears that 2c67 and 4v2 were uniquely expressed in lung.

Proteins detected in the highest number of intestine lysates were Cyps 2c29, 4a12a, and 3a25 (Figure 2-4E). Cyps 3a13, 2j5, 2c37, 4b1 and 4a10 were also detected. Cyps 4a12a and 4b1 were found in pregnant mice slightly higher than age-matched female lysates.

Cyp2c70 and 2j5 were the most frequently identified proteins in heart lysates (Figure 2-4F). These proteins were also expressed in highest abundance in the heart in comparison to other extrahepatic organs. Cyps 4a12a, 2f2, 3a25, and 4b1 were also detected. No sex specific, age specific, or pregnancy specific trends were identified in heart.

Activity assays using mouse liver microsomes and the drug efavirenz were performed to demonstrate differences in efavirenz metabolism between sex, age, and pregnancy status. First, we characterized efavirenz metabolism in mice. 8-hydroxyefavirenz (8-OHEFV), the prominent known metabolite formed in humans, was identified by matched retention time and matched spectrum of the 8-OHEFV standard (data not shown). A second species did not match the other known efavirenz metabolite, 7-hydroxyefavirenz, either in retention time or spectrum. We propose the identity of the novel compound is 5-hydroxyefavirenz (5-OHEFV). Figure 2-5 shows the MS2 spectrum and proposed fragment structures for the novel metabolite in the bottom panel. Fragments 210 and 246 match peaks in the 8-OHEFV spectrum. The structures for these fragments were previously published. Unpublished structures are proposed for the peak at 197 and base peak at 225, corresponding to fragmentation at various points in the carbon chain. In addition, we reasoned that if the aromatic ring can be hydroxylated in two locations (positions 7 and 8), the one position remaining on the ring (position 5) might similarly be available for hydroxylation. No significant formation of dihydroxylated metabolites was found in any of the conditions tested.

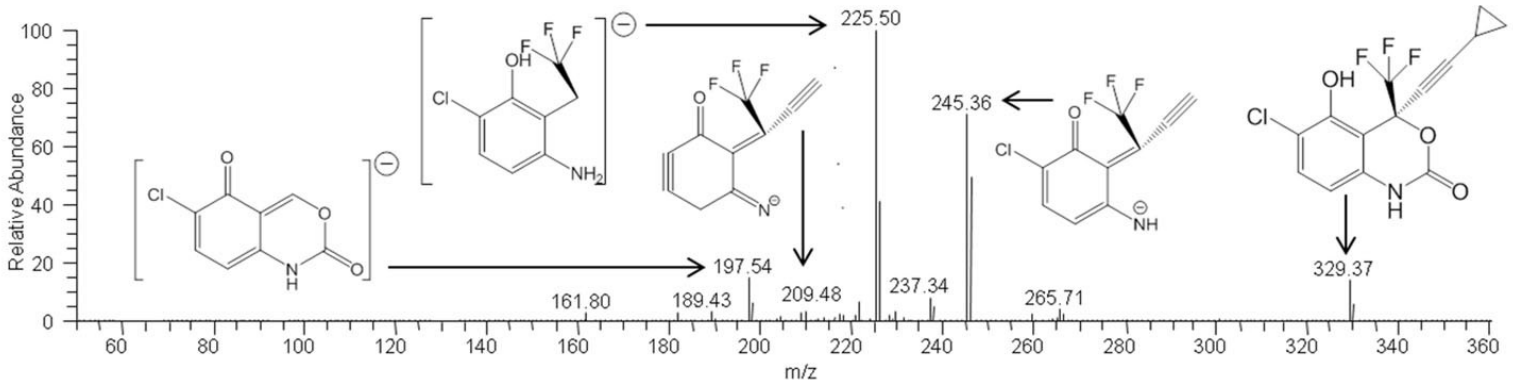


Figure 2-5. Fragmentation structure of 5-hydroxyefavirenz. The MS/MS spectrum of the peak eluting at 2.8 minutes is shown. The spectrum reflects a metabolite produced from a reaction with commercial Balb/c liver microsome enzymes (n=500). Proposed fragment structures of 5-OHEFV are indicated for all major peaks with arrows.

To determine the effect of age, sex, and pregnancy status on the metabolism of EFV to 8-OHEFV and 5-OHEFV, we compared the in vitro formation of metabolites by microsomes between groups. Areas under the chromatographic peaks (Figure 2-6A) were normalized to the F-EFV internal standard. We performed two-tailed unpaired t tests in GraphPad Prism; this revealed that 8OH-EFV (Figure. 2-6B) and 5-OHEFV (Figure. 2-6C) formation was reduced in 3-4 week old female microsomes compared to 9-10 week old female microsomes ($p=0.00353$ and $p=0.0392$, respectively), but not in male microsomes of similar ages. Similarly, 8-OHEFV (Figure. 2-6B) and 5-OHEFV (Figure. 2-6C) were formed less by 3-4 week-old female microsomes in comparison to 3 week-old male microsomes ($p=0.0359$ and $p=0.0102$, respectively), whereas there were no differences comparing young and older adult male and females microsome-generated metabolites. Microsomes from pregnant mice (10-12 weeks of age) demonstrated significantly higher 8-OHEFV (Figure. 2-6B) production in comparison to 9-10 week old females ($p=0.0153$), whereas 5-OHEFV (Figure. 2-6C) was significantly lower in pregnant mice microsome reactions in comparison to 9 week-old female microsomes ($p=0.0493$). Fetal microsomes demonstrated no activity against EFV (data not shown). These results suggest that metabolism of efavirenz differs with age, sex, and pregnancy status.

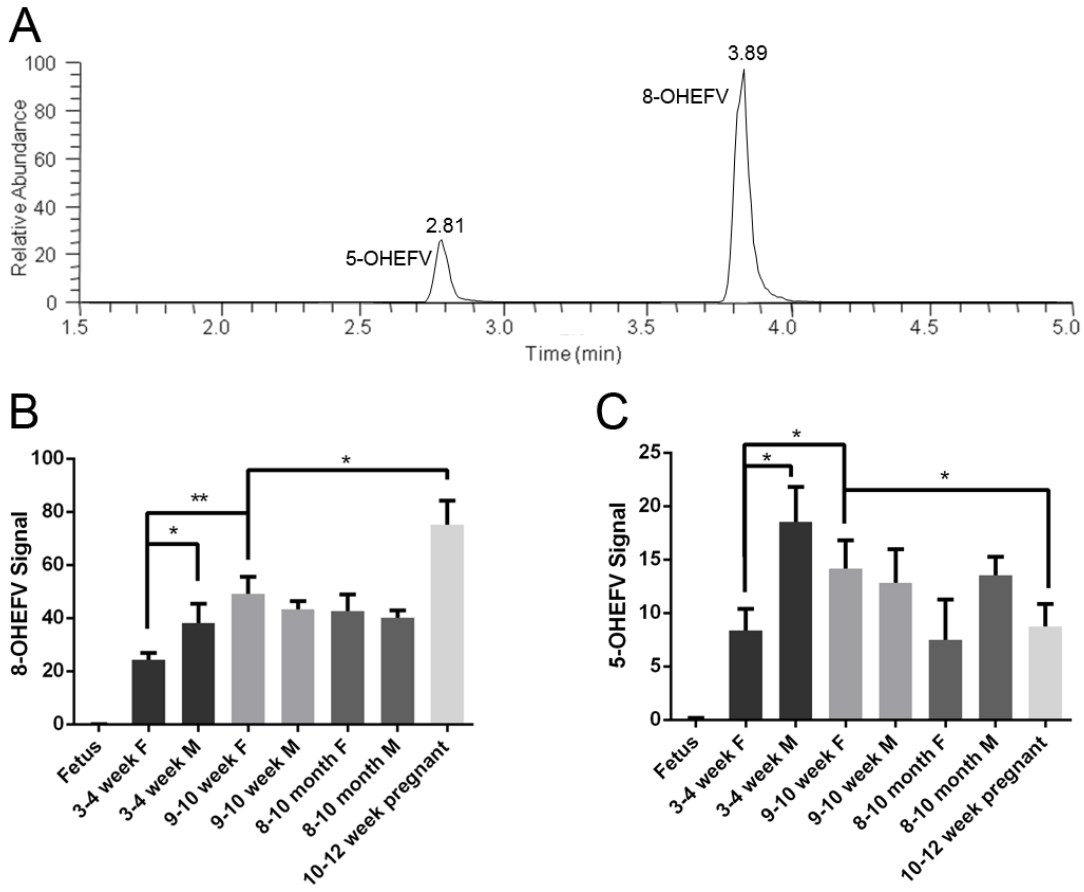

A
3.89
8-OHEFV
2.81
5-OHEFV
Relative Abundance
Time (min)
B
8-OHEFV Signal
Fetus
3-4 week F
3-4 week M
9-10 week F
9-10 week M
8-10 month F
8-10 month M
10-12 week pregnant
C
5-OHEFV Signal

Figure 2-6. Efavirenz monohydroxylated metabolite chromatograph and normalized peak areas comparing age, sex, and pregnancy status. A chromatograph of monohydroxylated efavirenz metabolites eluting from a C18 column, selecting for MRM parent/transition pair 329/117 is shown in (A). Metabolites were formed by a reaction with commercial Balb/c liver microsome enzymes (n=500). Differences in the formation of 8-hydroxyefavirenz (B) and 5-hydroxyefavirenz (C) are shown comparing age, sex and pregnancy status. Pooled microsomes (n=4) were prepared by liver homogenization and ultracentrifugation. Efavirenz was incubated with microsomes and NADPH for 30 minutes at 1 mg/ml. Metabolites were extracted using organic, and fluorinated efavirenz was used as internal standard for sample injection. Peak signal was calculated from MS2 mode chromatograph, and analyzed by T-test in GraphPad Prism.

Discussion

We have demonstrated the relationship between murine protein expression, and Cyp activity. We present the Cyp expression profile for mice of 8 metabolically relevant groups and 6 organ types, including 6 murine proteins never previously identified at the protein level. This is the first proteomic analysis of natural variation in cytochrome P450 biology.

Using global proteomic mass spectrometry, we identified 27 murine Cyps. Previous proteomic studies on murine Cyps identified a comparable number of proteins, although the isozymes identified differed between studies and mouse model systems (Flint et al, 2010; Sutton et al, 2010; Wright & Cravatt, 2007). Intensity of Cyp detection in liver microsomes is not indicative of expression in comparison to organ lysates due to the enrichment for ER membrane proteins. Comparison of tissue lysates between the two mass spectrometry methods revealed Cyps identified by the targeted method exceeded that of the global method by 4 in kidney, 11 in lung, 8 in intestine, 6 in heart and 4 in brain. Peptide detection depends on the complexity of the sample, the amount of protein analyzed, and the type of mass spectrometer among other factors. Absence of peptide detection may indicate either no protein, or levels below the limit of detection. The Cyps we identified include the murine isozymes structurally homologous to the main drug-metabolizing enzymes in human liver. These include 1a2, 2b9, 2c39, 3a11, and 2d9, which are structurally homologous to 1A2, 2B6, 2C19, 3A4, 2D6, and 3A4 respectively (Pham et al, 2011).

Next, we quantified 27 cytochrome P450s across 44 samples, describing protein distribution across tissue type, age, sex, and pregnancy status. Cyps 4b1, 3a25, 4a12a, and 2c29 were the most consistently expressed between mouse organs, with detectable amounts of protein identified in more than 5 of the 6 tissues tested. Results shown here demonstrate Cyp4b1 in all tissues, especially kidney and lung of ages. In particular Cyp4b1 expression was increased in adult male kidney, consistent with reports of dimorphic expression. Cyp1a2 expression appeared to be sexually dimorphic at all ages tested. Although this trend has not been substantiated by proteomics literature, male and female human hepatocytes have shown differences in CYP1A2 in response to inducers (Dhir et al, 2006). Cyp3a25 is not yet well characterized, mRNA has been shown to be expressed most prominently in liver and intestines similar to our proteomics results (Renaud et al, 2011). Cyp4a12a was well characterized by Muller et al 2007. They found that 4a12a contributed to eicosanoid metabolism more than the other 4a isozymes tested. Lower female Cyp4a12a mRNA and protein expression was found in comparison to males, which was an exclusively hepatic trend (Muller et al, 2007). Similar to the results presented here, Cyp2c29 mRNA is mainly expressed in liver and lung, but also in many other organs (Choudhary et al, 2003). Cyp2c29 detoxifies alpha- and beta-unsaturated aldehydes, found endogenously, in food, and in the environment, which can otherwise react with genomic material (Amunom et al, 2011). Most of the extrahepatic cytochromes P450 identified in this study demonstrate expression patterns consistent with the literatureEight Cyps were identified only in liver microsomes. Cyps 2c39 and

2c54 confirm previously reported mRNA expression hepatic profiles (Luo et al, 1998; Wang et al, 2004), whereas others have been reported in other tissues (Bieche et al, 2007; Renaud et al, 2011; Stiborova et al, 2013; Wang et al, 2004). The liver is the primary site for drug detoxification, therefore, some liver-specific expression was expected. Finally, the uncharacterized Cyp4f13 also demonstrated exclusively hepatic expression. Little is known about murine Cyp4f13, but the 4F subfamily of human and rats is involved in inflammation and drug metabolism and was identified by PCR to be expressed in kidney, lung and brain as well as liver (Kalsotra & Strobel, 2006).

Cyp2b9 is the most intensely expressed Cyp, reaching its highest levels in brain tissue. The enzyme functions in anti-carcinogenic activity (Pham et al, 2011) and steroid synthesis. The proteomics data presented here identified Cyp2b9 as expressed most highly in the brain compared to other Cyps, and no trends toward sexual dimorphism. Cyp2b9 protein was identified in fetal liver and brain samples in this study, but the mRNA primers for Cyp2b9 were not included in previous fetal hepatic mRNA screens (Choudhary et al, 2003; Choudhary et al, 2005). This result will require more investigation, although it could lead to an interesting difference between mRNA and protein abundances. In contrast to extrahepatic proteomic results presented here, Cyp2b9 mRNA expression was not detected in kidney, lung, intestine, heart or brain and was found in significantly higher abundance in female liver (Hernandez et al, 2006; Renaud et al, 2011; Xie et al, 2013). The detoxification roles of these enzymes might

explain their essential function, which requires body-wide expression. The cumulative extra-hepatic activity of these enzymes might contribute significantly to the total metabolism of an organism, which should be considered in experiments such as liver microsome activity assays.

Cyp4a10 is involved in salt-elimination from the body (Nakagawa et al, 2006). Because of this role, expression of Cyp4a10 was expected in the kidney, and, while it was not the most abundant Cyp, among all the extrahepatic organs where it was it expressed, kidney was the most abundant. No sexually dimorphic trends were observed in hypertension of the knockout mice (Nakagawa et al, 2006) in contrast with the mild trend for higher protein expression in 3-4 week and 9-10 week old female Cyp4a10 expression in kidney. Other Cyps expressed in the kidney also seemed to demonstrate sexual dimorphism in protein expression, including Cyps 2d26, 2b9, and 2e1. Cyp2d26 was highly expressed in all ages of female kidney lysates, without detection of the protein in either adult male tissue. Of these Cyps, no sexual dimorphism was detected in the corresponding liver microsomes. The trends for sexual dimorphism in adult kidney but not in liver are confirmed by mRNA expression (Dekant et al, 1995; Renaud et al, 2011). Conversely, it is known that mouse Cyp2e1 mRNA is higher expression in male tissues (Penaloza et al, 2013), confirming our finding of higher Cyp2e1 in kidney of all ages in comparison to female tissue lysates. Transgenic studies used male mice exclusively to investigate the biological mechanism for 2E1 in alcohol induced liver damage and high-fat mediated liver

damage (Abdelmegeed et al, 2012; Morgan et al, 2002). Results presented demonstrate the complexity of expression patterns, both hepatically and extrahepatically.

Cyp protein expression in lung also reflects expression patterns with human enzymes. In particular, Bieche et al. identified predominant expression of human CYP2F1 expression in lung in comparison to other extrahepatic tissues (Bieche et al, 2007). Similarly, Cyp2f2 was in the top 3 most abundant Cyps expressed in lung, and Renaud et al. identified comparable levels of Cyp2f2 expressed in lung in comparison to liver (Renaud et al, 2011). Cyps 2b9 and 4a12a were expressed equally high as Cyp2f2, although mRNA of Cyp 2b9 and 4a12a were not identified in lung (Renaud et al, 2011). Expression of other Cyps expressed in lung such as Cyps 2b10 and 2c70 were shown to be expressed in higher abundance in liver (Renaud et al, 2011) and we see evidence of these in some samples. Drug metabolizing enzymes in lung are particularly important for environmental toxins and drugs delivered by inhalation. Parkinson et al. are investigating bioactivation of a pulmonary toxin by Cyp4b1 using Cyp4b1-null mice (Parkinson et al, 2013). Murine models for drug metabolism inform researchers about the function and downstream pathways of a protein which could be complemented by mass spectrometric proteomic profiling to maximize the amount of information gained from the expensive transgenic experiments.

Intestine is often the first site of metabolism for oral medications, therefore it is one of the primary extrahepatic organs for study of xenobiotic metabolism. In

this study, the 3a and 2c subfamilies were identified in intestine, which are homologous to CYP3A and 2C subfamilies detected by immunoblot in human intestine (Ding & Kaminsky, 2003). The 4a, 4b, and 2j subfamilies were also identified in mouse intestine, which has not previously been shown in human or mouse intestine (Zhang et al, 2003).

Results demonstrate strong similarities between mouse and human cytochrome P450 subfamilies expressed in heart. Cyp2j5 was detected in mouse liver microsomes and mouse heart lysates abundantly in all mouse ages. Human CYP2J2 is well characterized in its localization to the heart (Bieche et al, 2007) as well as its xenobiotic and pathogenic role in heart disease (Xu et al, 2013). Interestingly, human subfamily CYP2C is known to contribute to ischemic heart disease, and while the mRNA has been shown to be not particularly abundant in heart (Renaud et al, 2011), the proteomic results presented in our work suggest strong expression of a murine Cyp2c family member, Cyp2c70, at all ages.

Knowledge of Cyp tissue distribution could be used to predict toxicity in a particular organ if it is a site for formation of a toxic metabolite. Similarly, knowledge of fetal Cyp expression can inform researchers about the mechanism of detoxification or potential formation of a toxic metabolite in neonates. Cyps 2b9, 3a25, and 4a12a were detected in fetal brain, which could have implications for local drug metabolism and potentially neuroprotection or toxicity which could affect neurodevelopment. The fetal liver expression data presented here, including Cyps 2d26, 2d10, 2b9, 2c37, and 2d9 represent a mutually exclusive

set of proteins in comparison to previous reports (Choudhary et al, 2003). Differences between these data and other published studies may be due to a number of differences in the experimental design, including the limit of detection of the peptides chosen to represent each murine Cyp in the targeted assay, but also due to failure in the assumption that mRNA expression correlates directly with protein expression.

Pregnancy in mice is known to induce or downregulate several Cyps (Koh et al, 2011). Our proteomics data demonstrates less expression of Cyp2e1 than other adult mice, a result contradictory to the work by Koh et al. Koh et al. did not test for expression of Cyp2b9, which was the only Cyp with increased expression in the data presented here.

Mouse Cyp activity against the HIV-therapeutic drug efavirenz revealed significant differences between mice of varying age, sex, and pregnancy status. Interestingly, the microsomes made from pregnant mouse livers formed more 8-OHEFV than age matched female mice. It has been shown that pregnant women taking efavirenz demonstrate increased metabolism of efavirenz in comparison to after birth (Cressey et al, 2012). It is interesting to note that 8-OHEFV and 5-OHEFV were significantly increased and decreased, respectively, indicating metabolism by multiple enzymes. In humans, multiple enzymes contribute to the metabolism of EFV, but CYP2B6 is known to contribute the most to the primary metabolite, 8-OHEFV (Ward et al, 2003). Cyp2b6 in humans increases from 9 days old to 28 years old (Tateishi et al, 1997). Cyp2b6

differences in age were significant between female mice of 3-4 weeks of age and the 9-10 weeks and 8-10 month old adults. Male mice did not exhibit a similar difference in activity against EFV with age, resulting in significant differences between male and female mice at the youngest age. Functional homology of Cyp2B6 has not been characterized for the mouse 2b enzymes. Proteomic profiles in this study demonstrate expression of 2b10 in liver microsomes and lung, without any striking trends in age-sex-or pregnancy related biology. Cyp2b9 is also homologous to CYP2B6, and does demonstrate some age- and pregnancy-related trends, although expression patterns do not correlate with activity patterns. Cyp2b9 was expressed in fetal liver microsomes, but no metabolite formation was detected in activity assays. The metabolism of EFV is known to be very complex in humans, and is likely to be similarly complex in mice, in which case expression patterns of a single enzyme may not be indicative of metabolite production. Further, establishing a causal relationship between murine Cyps and EFV metabolite formation would greatly advance an understanding of Cyp mechanistic biology and work towards a predictive model of drug metabolism.

Using targeted proteomics and activity assays, we have demonstrated a new experimental model for a more integrated approach to molecular pharmacology. These experiments are the first to describe protein expression in multiple tissues of male and female mice of multiple ages. Characterization of a mouse model for

the study of cytochrome CYP enzymes assists in a mechanistic understanding of drug metabolism.

Chapter 3:

Histone Post-translational Identification and Quantification using Isotope Labeled Derivatization

Introduction

The current approaches used in histone PTM analysis have not been able to elucidate site specific identification and quantification in a robust and accessible method. Bottom up mass spectrometry is ineffective in its application towards histones because histones have an abundance of lysine and arginine residues so that tryptic digestion yields many small peptides (2-4 residues). Small peptides are not easily observed and comprise too few and repetitive sequences that make it difficult to establish the location of the modification. Moreover, the principal sites for modification by acetylation, methylation, ubiquitylation, etc. are the lysine residues. When modified these sites generally become “missed cleavages” (Garcia et al, 2007b), producing different peptide segments for the modified and unmodified forms that are then difficult to compare quantitatively. Smaller fragments are also less likely to encompass multiple modification sites, and therefore do not provide quantification for combinatorial isoforms.

The top down approach addresses the peptide size and missed cleavage challenges in analyzing histones, but intact histone isoforms are relatively resistant to reverse-phase chromatographic separation (Siuti & Kelleher, 2007), and the targeting of individual isoforms for data-dependant MS/MS analysis becomes exceedingly difficult as multiple precursor ion species with similar retention times can result in missing the lower abundance isoforms. Hydrophilic interaction chromatography enhances histone application of the top down approach (Garcia et al, 2007c) and “middle-down” mass spectrometric analysis

using end protease Glu-C digestion has enabled focusing on the tail regions where the bulk of the modifications are found (Wang et al, 1999; Young et al, 2009). Both these methods have been successful, although they are cumbersome and often not site-specific.

The lysine deuteroacetylation technique is particularly advantageous for peptides containing multiple lysines because this approach can quantitate the numbers of lysines acetylated, distinguishing and quantifying the positional isomeric forms. When additional modifications, such as methylation, are also present in the same peptide, this approach can be used to quantitatively compare those isoforms differing in acetylation but having a common methylation, or other modification, site. The deuteroacetyl label eliminates several “biases” common in quantitation methods, since enzymatic digestion, chromatographic separation, and ionization efficiency (in the mass spectrometer) are effectively normalized for all analyzed samples. In our laboratory we have used deuteroacetylation with MALDI tandem time-of-flight mass spectrometry previously to determined the percent acetylation at K56 on histone $H3_{54\text{-}63}$ (FQK_{56}STELLIR) in yeast deficient in sirtuins Hst3 and Hst4p (Celic et al, 2006) and identified the major isoforms spanning residues K5, K8, K12, and K16 in yeast histone H4 (Cotter et al, 2007).

In this context, we have continued to develop a bottom up approach that exploits and extends the advantages of prior lysine deuteroacetylation using high performance mass spectrometry and the application of the approach to other modifications. Specifically, both MALDI TOF and high resolution

nanospray/Orbitrap mass spectrometry are used here to characterize the positional microheterogeneity of the histone H4 tail region acetylation isoforms in HeLa cells and to observe the specific changes to the H4 tail region for cells treated with trichostatin A and nicotinamide, two broad spectrum inhibitors of histone deacetylases. This model system is used to describe the approach developed to quantify acetylations. In addition, deuteroacetylation enables the observation of methylated (lysine and arginine) species in the context of larger peptides that may also be acetylated. In several examples reported here it is shown that monomethylated lysines are derivatized by deuteroacetylation, while dimethylated and trimethylated lysines are not. While the ability to distinguish isobars with the same nominal mass is accomplished relatively quantitatively for methylated as well as acetylated species using this approach, the additional advantages of high mass resolution and accuracy to distinguish acetyl and trimethyl modifications are also shown.

Methods

Preparation of yeast histones. Yeast cells containing the appropriate plasmids were grown as reported previously (Celic et al, 2006). Briefly, yeast cells were grown at 30^{o}C in YPD medium and arrested in G1 phase using 10 μg/ml α-factor for 2-3 hours. The cells were harvested directly or released into the cell cycle in fresh YPD medium. Harvested cells were lysed, nuclei extracted, and histone proteins were acid precipitated prior to SDS PAGE to resolve individual histones.

HeLa cell culture and histone purification. HeLa cells were cultured in Dulbecco's modified eagle's medium containing 10% fetal bovine sera (GIBCO; Invitrogen, Carlsbad, CA). Cells were plated and grown to 90% confluence. One sample was treated with deacetylase inhibitors: 1 μM Trichostatin A (TSA, Sigma-Aldrich, St. Louis, MO) and 10 mM nicotinamide (NIA, Sigma-Aldrich, St. Louis, MO) for three hours while the other sample was supplemented with an equal volume of solvent alone. TSA and NIA were dissolved in DMSO and deionized water, respectively. Cells were then harvested, pelleted and fractionated to isolate nuclei. Histones were acid extracted as described in the method by Wang et al. (Wang et al, 1999) that provides relatively purified histones, and resolved by SDS-PAGE (4-12% NuPAGE™ gels; Invitrogen, Carlsbad, CA) with purified chicken core histones (Millipore; Billerica, MA) run as markers, and the gel was stained with Coomassie Brilliant Blue.

Deutero (d3)-acetylation, chemical derivatization and digestion. Using reagents provided in the Trypsin Profiler IGD Kit for in-gel digestions (Sigma-Aldrich, St. Louis, MO), CBB-stained HeLa H4 gel bands were excised, cut into 1 mm pieces and destained for one hour with a 200 mM ammonium bicarbonate solution containing 40% acetonitrile. The destained protein bands were resuspended in a solution containing 50 μL deuterated (d_4) acetic acid and 10 μL deuterated (d_6) acetic anhydride (deuterated reagents; Sigma-Aldrich, St. Louis, MO). To ensure the derivatization reaction went to completion, gel bands were incubated in the deuterated reagents for five hours at room temperature. The treated gel slices

were then rinsed with distilled water, titrated with 200 mM ammonium bicarbonate (Sigma-Aldrich, St. Louis, MO) to pH 8, rinsed with water again to remove remaining buffer salts, and dehydrated/dried in a Speed Vac. Proteins were digested in-gel with 0.4 µg trypsin (1 mg/ml in 1 mM HCl) in 40 mM ammonium bicarbonate containing 9% acetonitrile, incubated overnight at 37°C (Jimenez et al, 2001). Resultant peptides were extracted from the gel and lyophilized to dryness before mass spectrometric analysis. All samples were reconstituted as stock solutions in 5 µL water with 0.1% trifluoroacetic acid (Pierce, Rockford, IL).

MALDI-TOF mass spectrometry. Matrix solution was prepared as 10 mg/mL α-cyano-4-hydroxycinnamic acid (Sigma Aldrich, St. Louis, MO) dissolved in a 1:1 solution of acetonitrile:deionized water containing 0.1% trifluoroacetic acid. Undissolved matrix particles were removed via centrifugation. Using the dried droplet application technique, resultant matrix solution (1 µL) and the resuspended histone H3 and H4 peptide mixtures (1 µL) were spotted on the stainless steel MALDI target (Cohen & Chait, 1996). To ensure co-crystallization, target plates were spotted at room temperature using the sandwich method: deposit 1µL of matrix, 1µL of digest and 1µL of matrix, then redissolve in 50% acetonitrile containing 0.1% trifluoroacetic acid.

Deuteroacetylated HeLa histone H4 peptides were analyzed using a Shimadzu AXIMA-TOF^2 tandem time-of-flight mass spectrometer (Manchester, UK) equipped with a 337 nm pulsed nitrogen laser, high energy collision chamber and

curved-field reflectron (Cornish & Cotter, 1993; Cotter et al, 2004). The acceleration voltage was set at 20 kV. A four point, external calibration was applied per sample spot using a peptide mixture containing bradykinin (*m/z* 757.40), angiotensin II (*m/z* 1046.54), $P_{14}R$ (*m/z* 1533.86), adrenocorticotropic hormone fragment 18-39 (*m/z* 2465.20), and insulin B chain (*m/z* 3495.65). All peptides used in the calibration mixture were from Sigma Aldrich (St. Louis, MO). Each mass spectrum acquired consisted of an average of 500 profiles of 10 shots accumulated per profile.

Sample preparation for HPLC-ESI-MS/MS Analysis. 4 µL of stock deuteroacetylated peptide mixture were diluted 2.5 fold (to a final volume of 10 µL) and placed in an Agilent 1200 autosampler (Agilent, Santa Clara, CA). 5 µL of each 10 µL sample were loaded by the autosampler through a trapping column onto a fused silica PicoFrit (New Objective, Woburn, MA) capillary column, 75 µm inner diameter (i.d.) x120 mm long column packed in-house with 5 µm, 300 Å BioBasic C_{18} (Thermo Electron, Bremen, Germany) stationary phase, at a flow rate of 300 µl/min. Peptides were separated on-line via reversed phase nano HPLC using the Eksigent Nano 2D HPLC pumping system (Eksigent, Dublin, CA). The Eksigent Nano 2D HPLC system was controlled by XCalibur software, Version 2.0 (Thermo Electron, San Jose, CA). Separations were performed at mobile phase flow rate of 300 nL/min on the binary pump system using 0.1 % formic acid in deionized water (Solvent A) and 90% (vol/vol) acetonitrile with 0.1% (vol/vol) formic acid (Solvent B) using a linear gradient of 20-50% A over 60

minutes. The outlet flow of the nano-HPLC interfaced directly with the inlet of an LTQ-Orbitrap-XL (Thermo Electron, San Jose, CA) allowing for introduction of the analyte into the mass spectrometer.

Tandem Mass Spectrometric Analysis (LTQ Orbitrap XL). The LTQ-OrbitrapXL mass spectrometer was operated in data dependent mode. MS precursor scan spectra (*m/z* 300-2000) were acquired in the Orbitrap with mass resolution of 60,000; the six most intense ions from each MS scan were automatically targeted for fragmentation (MS/MS) in the ion trap. Using nitrogen as the collision gas, CID mediated peptide fragmentation in the linear ion trap. The LTQ OrbitrapXL was controlled by XCalibur software. For each experiment, the source voltage was set at 2.4 V, the capillary voltage at 48 V, and the capillary temperature at 200˚C. Sheath and auxiliary gases were not necessary because nanoflow parameters were being used. The tube lens voltage was kept at 105 V, and the ion gauge pressure was 1.5×10^{-5} Torr. The automatic gain control was used at the manufacturers default settings for MS collection in the Orbitrap (200,000) and MS/MS collection in the ion trap (10,000). Normalized collision energy was established at 35% for MS/MS. The default charge state was set at two. The isolation window for precursor ion selection was fixed at two Daltons. After an initial MS/MS fragmentation event, ions were excluded from additional fragmentation rounds for 30 seconds using dynamic exclusion. The ion selection threshold, the minimum signal required to trigger tandem mass spectrometry, was set to 500. The activation Q was set at 0.25.

Quantitation of peptide isoforms. Determining the relative ration between acetyl (d_0) and deuteroacetyl (d_3) ion pairs gives the relative abundance of the in vivo acetylated and unmodified peptide forms. The relative acetylation of each peptide was determined by an in-house processing method that compares the peak height of acetylated peptides to the total for all observed forms of the peptide as measured from the charge-state reduced chromatograph of the peptides' elution peak (generated using Xtract software package within XCalibur). Charge state reduction was employed such that the peak intensities from all charge states of each species were summed over the elution profile. Determining the relative ratio between acetyl (d_0) and deuteroacetyl (d_3) ion pairs gives the relative abundance between in vivo acetylated and unmodified peptide forms. To determine the relative abundance of positional isomers, where all isomeric forms fragment from the same precursor mass in the same MS/MS scan, we compared the peak intensities from selected b and y ion pairs distinguishing these isoforms (Garcia et al, 2005). Peak intensities from b and y ion pairs were summed together and this value was then divided by the total produced between both species.

Results

Acetyl Quantitation. Figure 3-1 shows the MALDI time-of-flight mass spectrum of the tryptic digest of deuteroacetylated histone H3 from the *hst3–H184A* yeast mutant, a deletion mutant of the Sirtuin Hst3. The peptide FQKSTELLIR encompasses the K56 residue, an acetylation site just inside the histone core

region. In the expanded mass spectrum (Figure 3-1, inset) the acetylated $FQK_{Ac}STELLIR$ form is observed at *m/z* 1276.94 along with its molecular ion distribution resulting from naturally occurring isotopes. Without derivatization, the unacetylated analog could not be observed, as K_{56} would be a tryptic cleavage site. With derivatization, the unacetylated analog is also observed as the deuteroacetylated peptide $FQK_{dAc}STELLIR$ three mass units higher at *m/z* 1279.95. Quantitative assessment of the degree of acetylation at this site is then based upon the fact that both species are chemically equivalent and that the ionization efficiencies are also equivalent. In this case 43% acetylation is determined after correction for the contribution of the 3rd isotope peak in the isotopic distribution of the naturally acetylated species to the peak intensity of the monoisotopic peak of the deuteroacetylated species. To determine the distribution in acetylation, the isotopic distributions of the endogenous and deuteroacetylated H4 peptides were simulated based on calculations using MS-isotope (Protein Prospector Software), available online at http://prospector.ucsf.edu/. For this mass and elemental composition the 3rd isotopic peak in the naturally acetylated peptide will have an area of 8.60% of that of its monoisotopic mass (Figure 3-2).

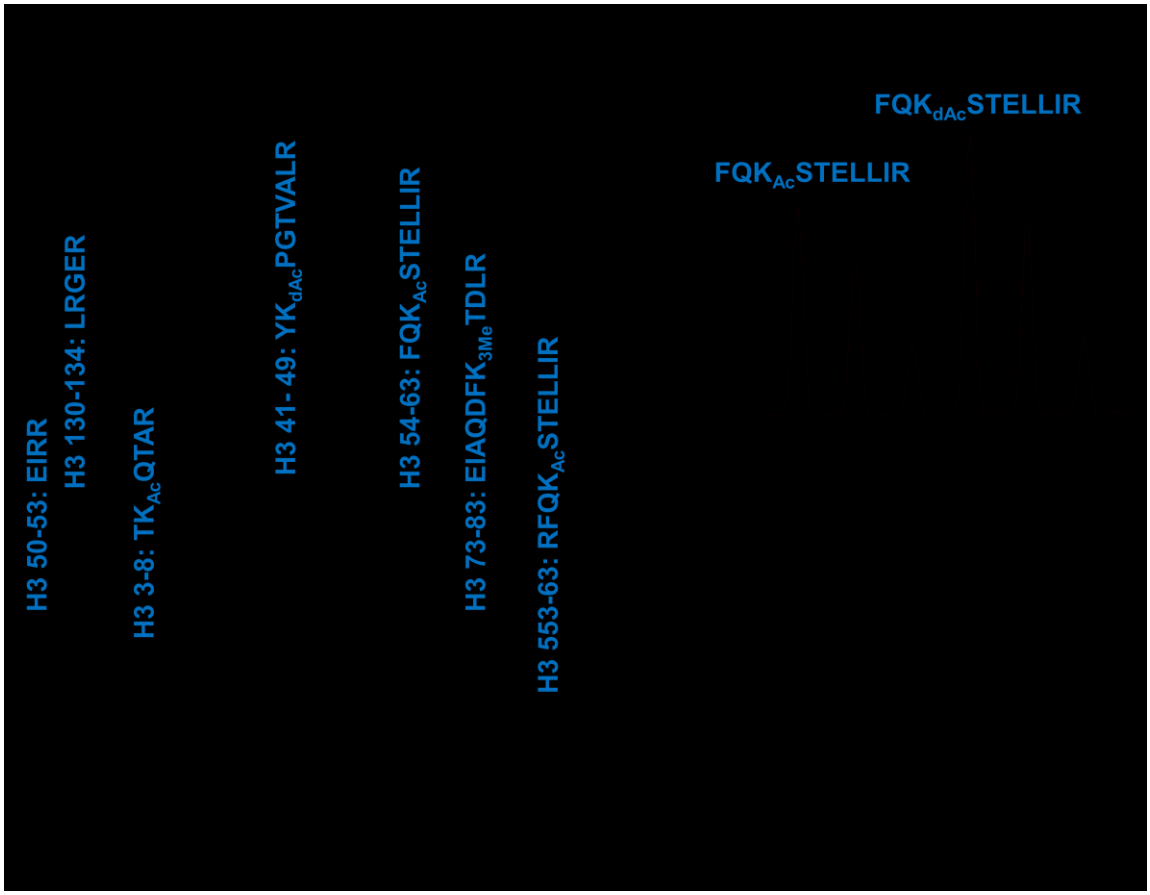


Figure 3-1. MALDI TOF mass spectrum of the tryptic digest of deuteroacetylated histone H3. Masses corresponding to the peptides expected from several of the known post-translational modifications are annotated. Inset: expansion of the H3 54-63 peptide showing the distribution of naturally acetylated and deuteroacetylated species.

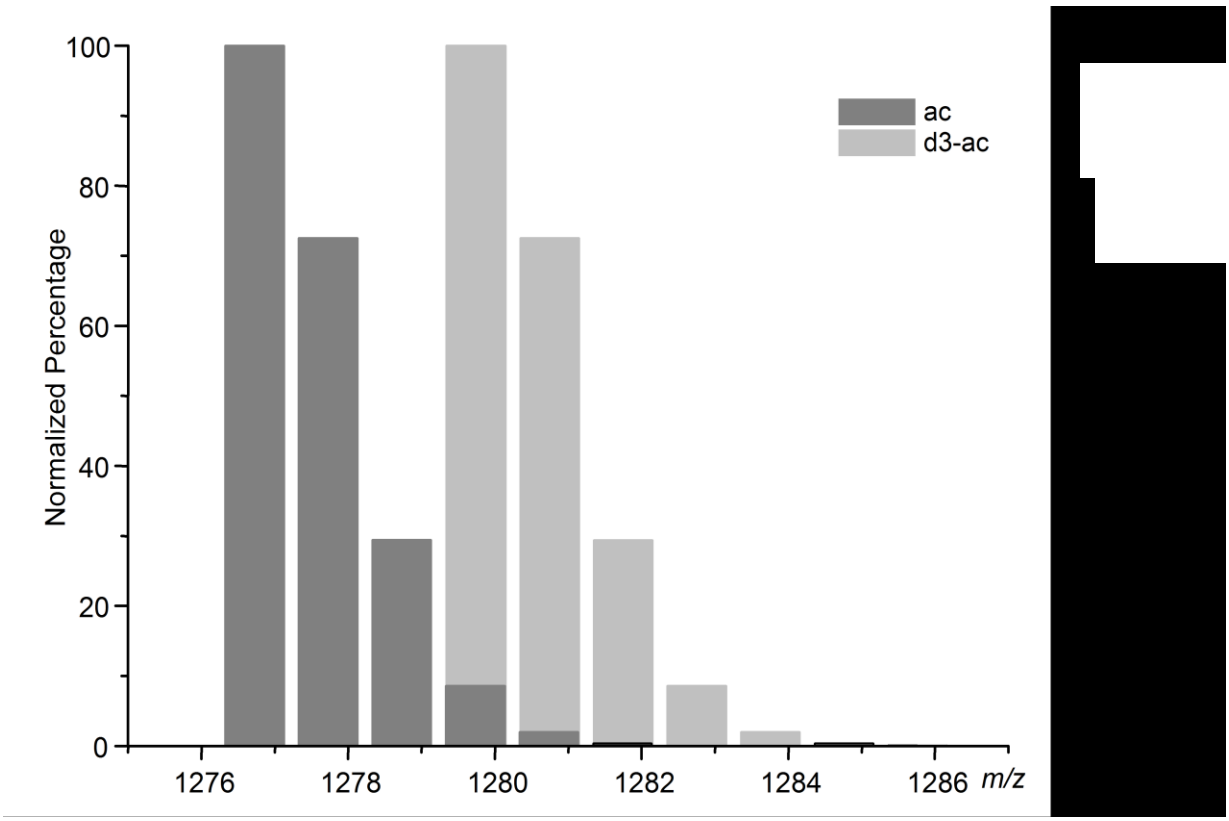


Figure 3-2: An overlaid plot of two isotopic distributions. Note the overlap of the 4th isotope peaks of the naturally acetylated peptide with the monoisotopic peak of the d3-acetylated peptide. This overlap accounts for 8.60% of the area of the monoisotopic peak of the naturally acetylated peptide.

Acetylated H4 tail peptide $GK_5GGK_8GLGK_{12}GGAK_{16}R$. Figure 3-3 is the MALDI TOF mass spectrum of the tryptic digest of deuteroacetylated histone H4 from HeLa cells, again showing the most prominent expected peptides. Of particular interest is the histone H4 tail peptide GKGGKGLGKGGAKR encompassing residues 4-17.

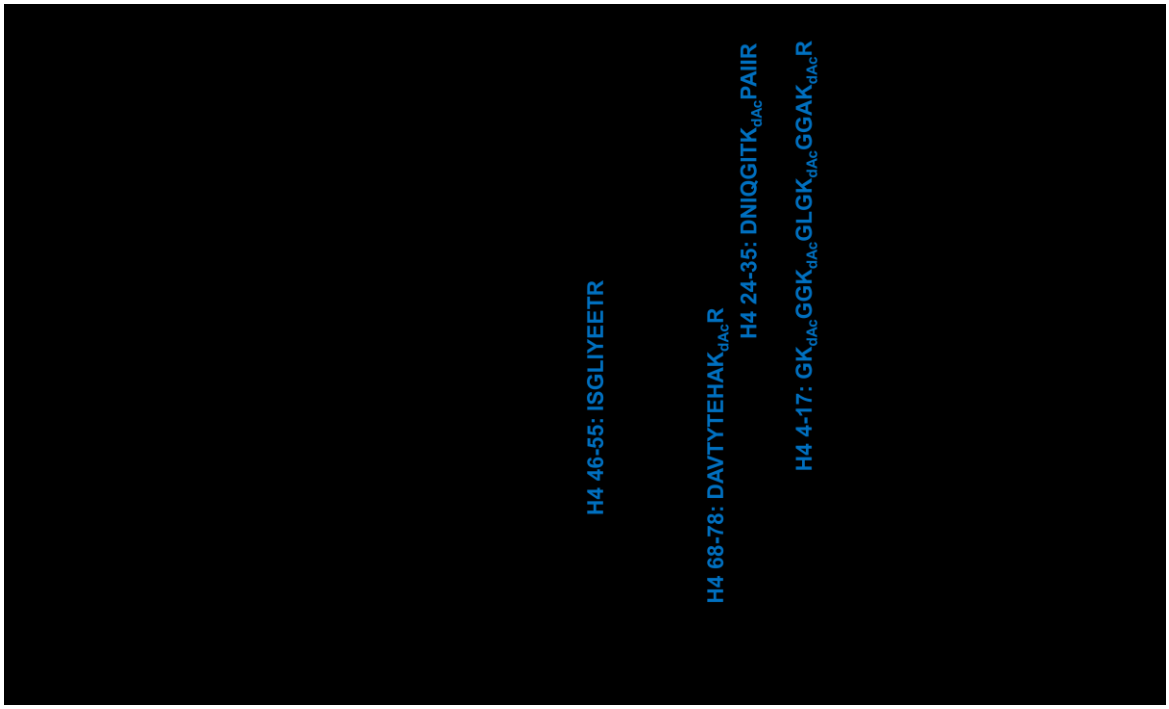


Figure 3-3. MALDI TOF mass spectrum of the tryptic digest of purified and deuteroacetylated histone H4. Masses corresponding to the peptides expected from several of the known post-translational modifications are annotated.

In our laboratory, this same approach was used on multiple species of histones, including yeast histones and human peripheral blood mononuclear cells (Figure 3-4). We demonstrate using MALDI TOF analysis that deuteroacetylation is effective with a variety of histone samples isolated and prepared from multiple methods and species.

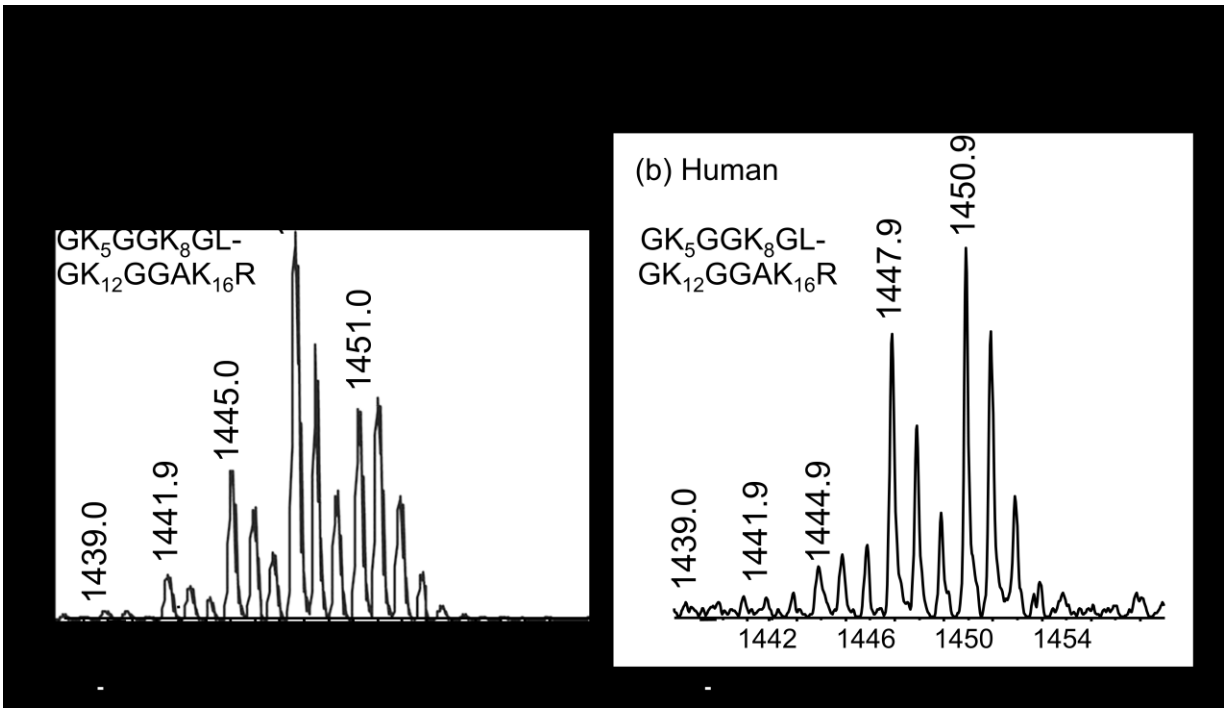


Figure 3-4. Endogenous acetylation distribution on N-terminal H4 peptide GKGGKGLGKGGAKR between species. Histones were isolated from (a) genetically modified his-tagged histone H4 yeast growing in log phase and (b) human peripheral blood mononuclear cells. Yeast histones were collected by his-tag purification in the lab of Jef Boeke and peripheral blood mononuclear cells were isolated by acid extraction from blood of healthy volunteers in the lab of Dr. Craig Hendrix. Spectra are labeled according to non-acetylated (d_{12}) to tetra-acetylated (d_0) species.

In our laboratory, this same approach was used with both MALDI TOF and ESI Orbitrap mass spectrometry to monitor effects of the histone deacetylase enzyme inhibitors trichostatin A and nicotinamide on the extent of acetylation and on the specific distribution of the positional isomeric forms. The MALDI mass spectra in Figure 3-5 was used to determine the number of acetylated lysine residues for the peptide GKGGKGLGKGGAKR obtained from HeLa cells treated with trichostatin A and nicotinamide (Figure. 3-5a) or control (Figure. 3-5b). Following derivatization all lysine residues are either acetylated (+42 Da) or deuteroacetylated (+45 Da), so that the lowest mass d_0 at *m/z* 1439.03 in this isotopic distribution corresponds to GKGGKGLGKGGAKR that is fully acetylated. A single acetylation corresponds to the d_9 peak, containing one acetylated lysine and three deuteroacetylated lysines. The unacetylated peptide is the d_{12} peak at *m/z* 1451.09. All of the peptides are chemically equivalent and the distribution is obtained from the peak heights of the d_0, d_3, d_6, d_9 and d_{12} species. Then, when calculating the actual abundance of each differently modified peptide, the contribution of the 3rd isotope peak of the lighter isoforms to the intensity of the monoisotopic peak of the isoform heavier by 3 Da must be subtracted. For example, the corrected intensity of the d_3 $GK_5GGK_8GLGK_{12}GGAK_{16}R$ peptide species is calculated by subtracting 10.75% of the d_0 monoisotopic peak from the measured d_3 peak intensity (Figure 3-6). Corrected distributions for the unacetylated to tetra-acetylated species obtained in the MALDI TOF mass spectra are shown in Table 3-1 for both the endogenous and deacetylase inhibited HeLa cell samples.

Figure 3-5b shows the analogous distribution for the histone tail peptide from untreated HeLa cells. Peptides with four acetylated lysines are not observed, and in fact three acetylated lysines are of very low abundance. Most of the peptides are unacetylated or mono-acetylated. Comparison with Figure 3-5a shows the effect of the histone deacetylase enzyme inhibitors in promoting a high degree of acetylation. From Table 3-1, approximately 57% of isomers present in the endogenous HeLa H4 tail peptide had no acetylated lysine, i.e. the most abundant isoform of this peptide was unmodified. 31% of the peptide had one acetylated lysine, and 9% had two acetylated lysine residues. Less than four percent of the $GK_5GGK_8GLGK_{12}GGAK_{16}R$ isoforms contained three or four acetylated lysines.

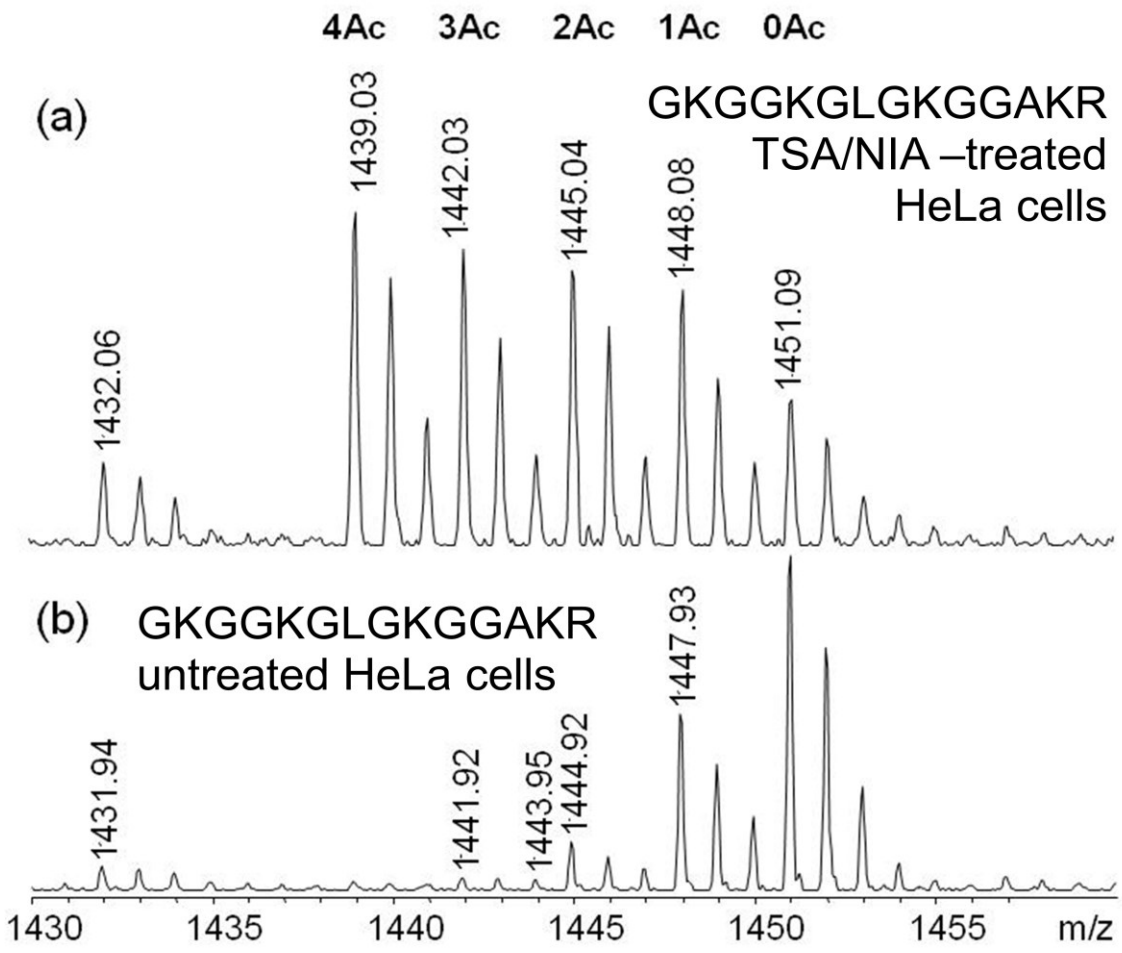


Figure 3-5. MALDI TOF mass spectra of the expansion of GKGGKGLGKGGAKR for HeLa cells treated with deacetylase inhibitor and control. The distribution of the non-acetylated (d_{12}) to tetra-acetylated (d_0) species is shown for (a) histone derived from trichostatin A/nicotinamide (TSA/NIA)-treated HeLa cells and (b) histone from untreated HeLa cells.

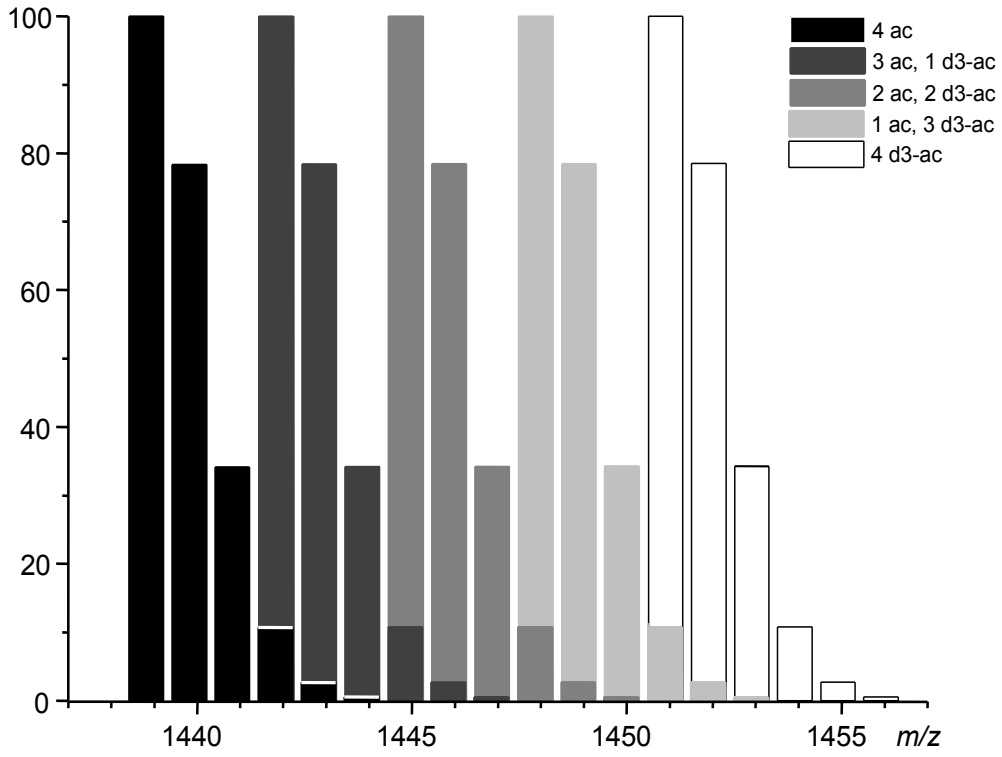


Figure 3-6. Calculated Isotope distributions for various isoforms of the deuteroacetylated GKGGKGLGKGGAKR. Histone H4 tail tryptic peptide (residues 4-17) S2F shows an overlaid plot of the isotopic distributions of all isoforms. Note the overlap of the 4 isotope peaks with monoisotopic peak of the isoform 3 Da heavier. This overlap is ca. 10.75% of the area of the monoisotopic peak.

Table 3-1. Relative abundances of each differentially modified peptide population for the model GKGGKGLGKGGAKR peptide. A comparison of isotopically corrected data derived from both the LC-ESI MS and MALDI-TOF is shown.

	No deacetylase inhibitor		With deacetylase inhibitors Trichostatin and Nicotinamide	
	MALDI TOF	ESI Orbitrap	MALDI TOF	ESI Orbitrap
No acetylation	57%	63%	10%	8%
1 acetylation	31%	29%	19%	14%
2 acetylations	9%	7%	21%	20%
3 acetylations	3%	1%	22%	26%
4 acetylations	0%	0.3%	28%	32%

Quantitation of H4 tail peptide isoforms using ESI Orbitrap MS/MS analysis. Tryptic digests of the deuteroacetylated histones were analyzed using on-line HPLC MS/MS. We confirmed that all 16 of the acetylated/deuteroacetylated chemically and chromatographically equivalent species of the tail peptide $GK_5GGK_8GLGK_{12}GGAK_{16}R$ eluted from the column at the same time (Figure 3-7), including the fully deuteroacetylated (b), singly acetylated (c), doubly acetylated (d), triply acetylated (e), and quadruply acetylated (f). While some small chromatographic differences were detected for peptides with different numbers of in vivo acetylations (note the peak spread in Figure 3-7a), by contrast positional isoforms, with the same number of naturally-occurring acetylations (and same molecular weights), had identical narrow retention times. Figure 3-8 is the ESI Orbitrap mass spectrum of the doubly-charged molecular ions of the tail peptides from the (a) untreated and (b) trichostatin A/nicotinamide treated HeLa samples. The extracted (isotopically corrected) relative abundances of the unacetylated to tetra-acetylated species are shown in Table 3-1, and are similar to those obtained by MALDI.

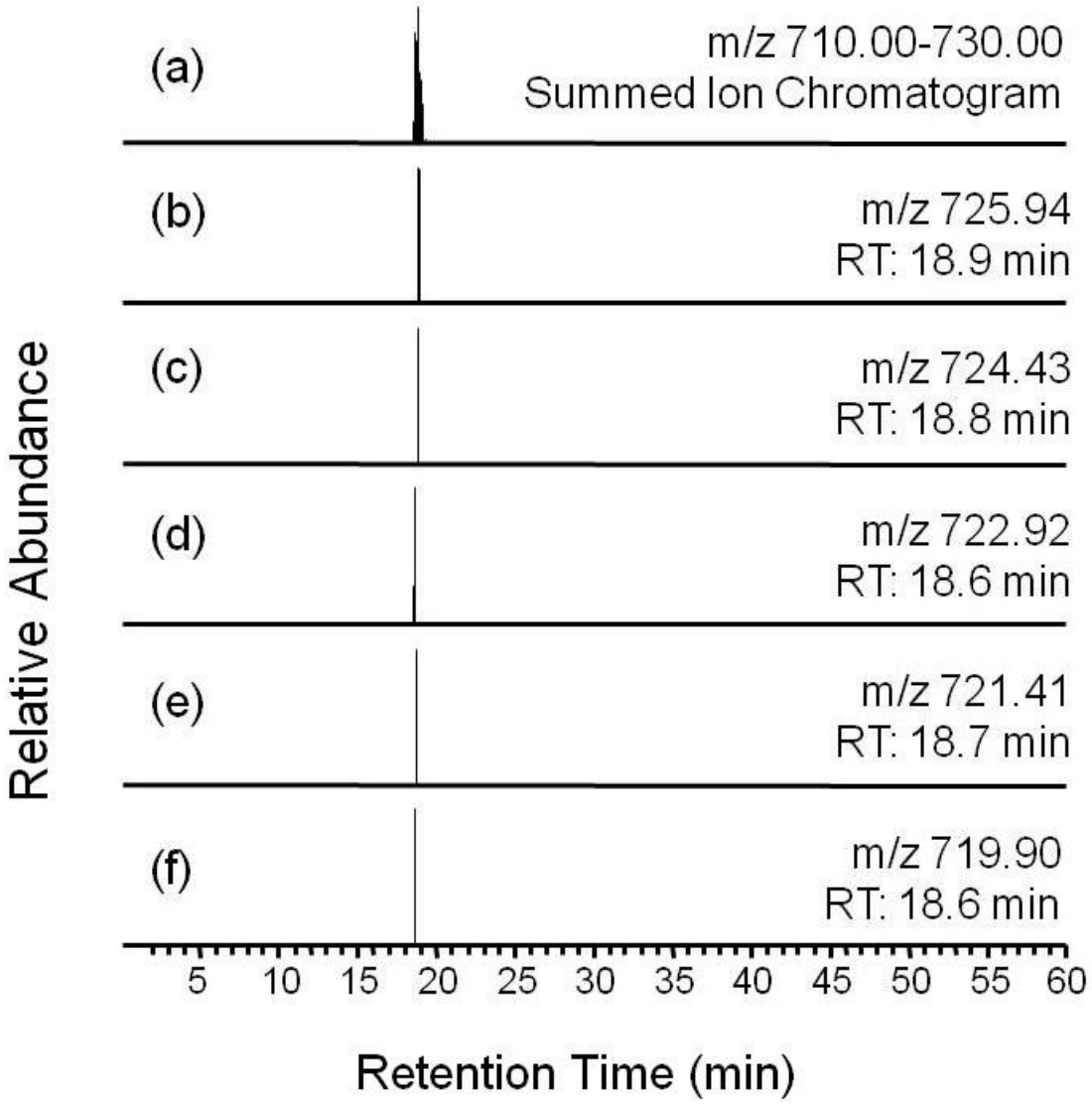


Figure 3-7. Chromatographic traces for the acetylated/deuteroacetylated isoforms of GKGGKGLGKGGAKR. The H4 4-17 peptide GKGGKGLGKGGAKR doubly-charged molecular ions elution from a C18 column interfaced to a nanospray ionization source were monitored on an LTQ/Orbitrap mass spectrometer.

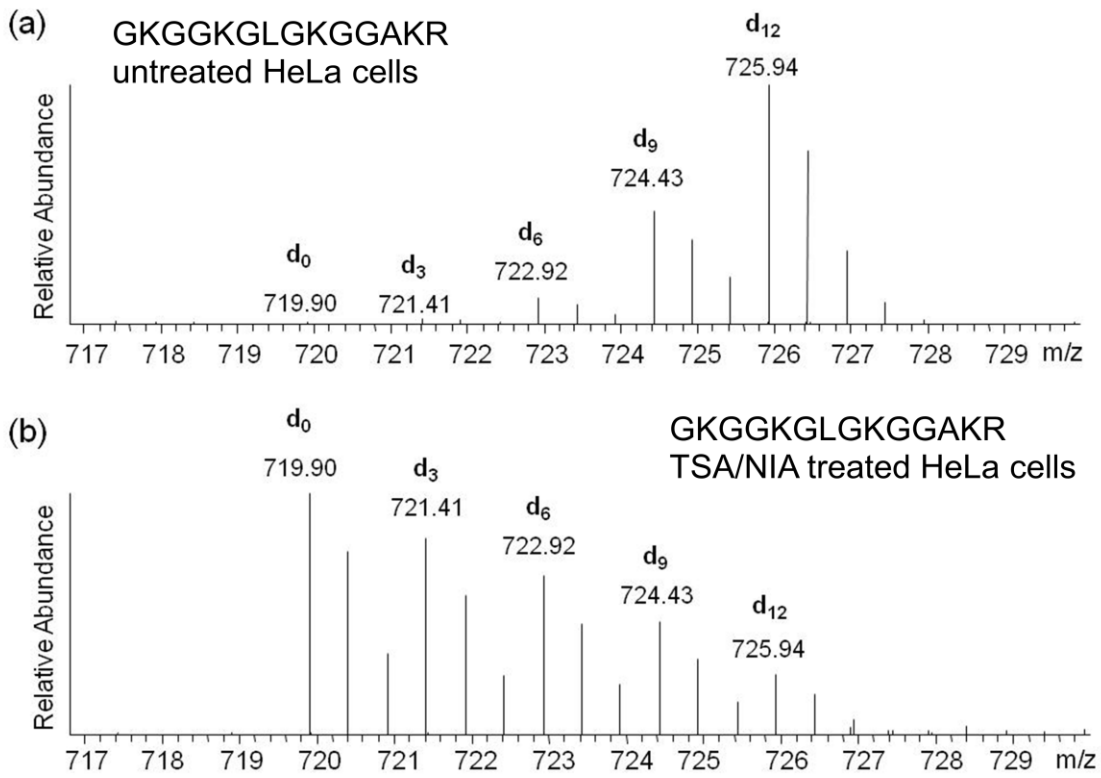


Figure 3-8. ESI Orbitrap mass spectra of GKGGKGLGKGGAKR for HeLa cells treated with deacetylase inhibitor and control. The doubly-charged molecular ion regions for the acetylated/deuteroacetylated isoforms of peptide GKGGKGLGKGGAKR integrated across the chromatographic retention times are shown in (a) untreated HeLa cells and (b) TSA/NIA-treated HeLa cells.

Doubly-charged molecular ions of the unacetylated to tetra-acetylated tail peptides $GK_5GGK_8GLGK_{12}GGAK_{16}R$ were each targeted for MS/MS fragmentation and the tandem mass spectra from the elution peak were summed. Figure 3-9 shows the MS/MS spectrum of the doubly-charged molecular ion *m/z* 719.90, which corresponds to the d_0 (fully-acetylated) species from the TSA/NIA-treated sample. There is only one positional isomer with this molecular weight giving rise to a single set of b-series and y-series ions. From these we selected three b and y fragment ion "pairs" that describe the fragmentation between the four lysine residues:

In choosing these ions we selected those that appear with reasonable intensity in all of the subsequent spectra of mixed isoforms. The standard deviation between selection of b in comparison to y ions according to cleavage at the same bond (e.g. y5 and b9, or y7 and b5) was determined to be less than 5%.

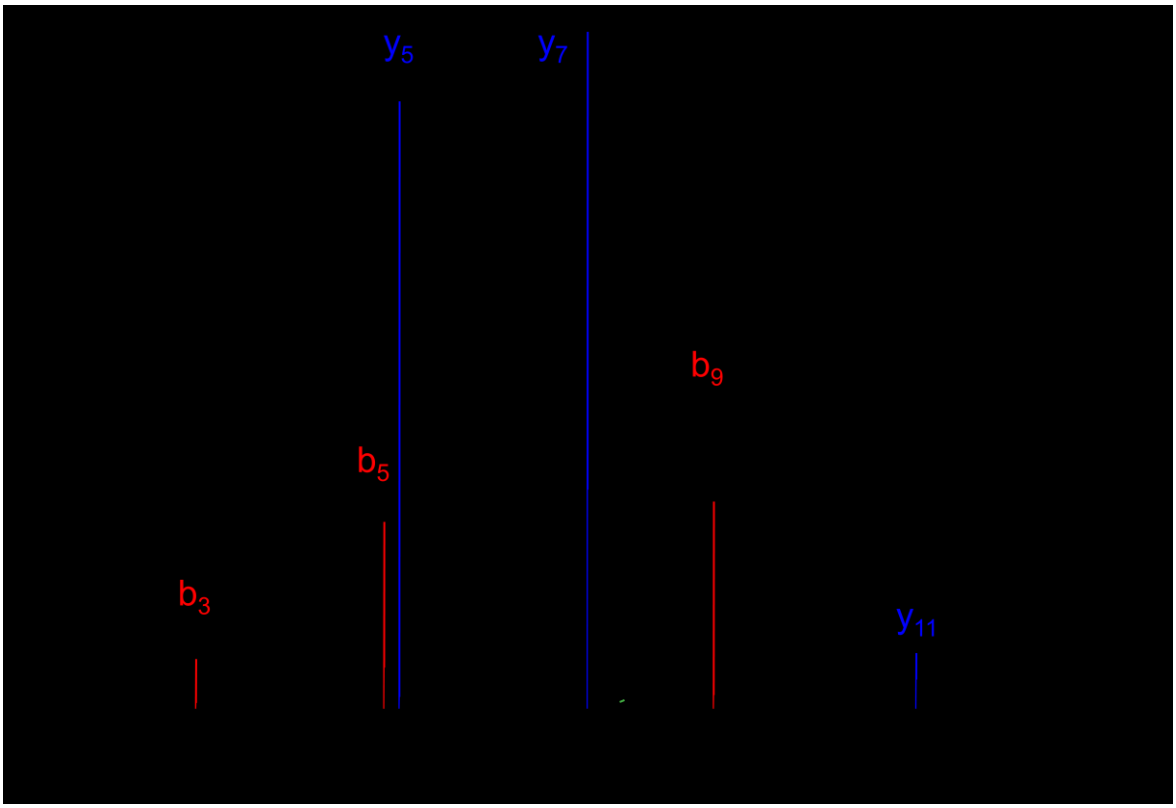


Figure 3-9. ESI Orbitrap MS/MS spectrum of fully acetylated GKGGKGLGKGGAKR. The doubly-charged molecular ion of the fully acetylated (d_0) isoform of the peptide GKGGKGLGKGGAKR is shown highlighting the b-series (b_3, b_5 and b_9) and y-series (y_5, y_7 and y_{11}) ions used to quantitate the positional isomers.

Figure 3-10 shows expansions of the mass regions about the six fragment ion masses for the mono-acetylated isomers. Each of these b and y ions appears at two masses according to an acetylated and a deuteroacetylated isoform and can be used to describe the relative abundances of the isomeric forms. For example: b_3 shows the ratio between ions acetylated at K_5 and those acetylated at any of the other three lysines. The y_{11} fragment does the same, though the order of mass and acetylation/deuteroacetylation assignment is reversed because the N-/C-terminal end of the peptide (b_3 or y_{11}) is reversed. These two ratios were averaged to produce the entry on Table 3-2 of 7.3% acetylation at K_5 and 93% acetylation on all other lysines. The b_5 and y_7 ratios then determine the relative abundances of acetylation at K_5 or K_8 versus K_{12} or K_{16}. And the b_9 and y_5 ratios determine the percentage acetylated at K_{16}. Combining these data (Table 3-2), we determined that the composition for the mono-acetylated isomers from untreated cells is 7% K_5, 5% K_8, 13% K_{12} and 74% K_{16}.

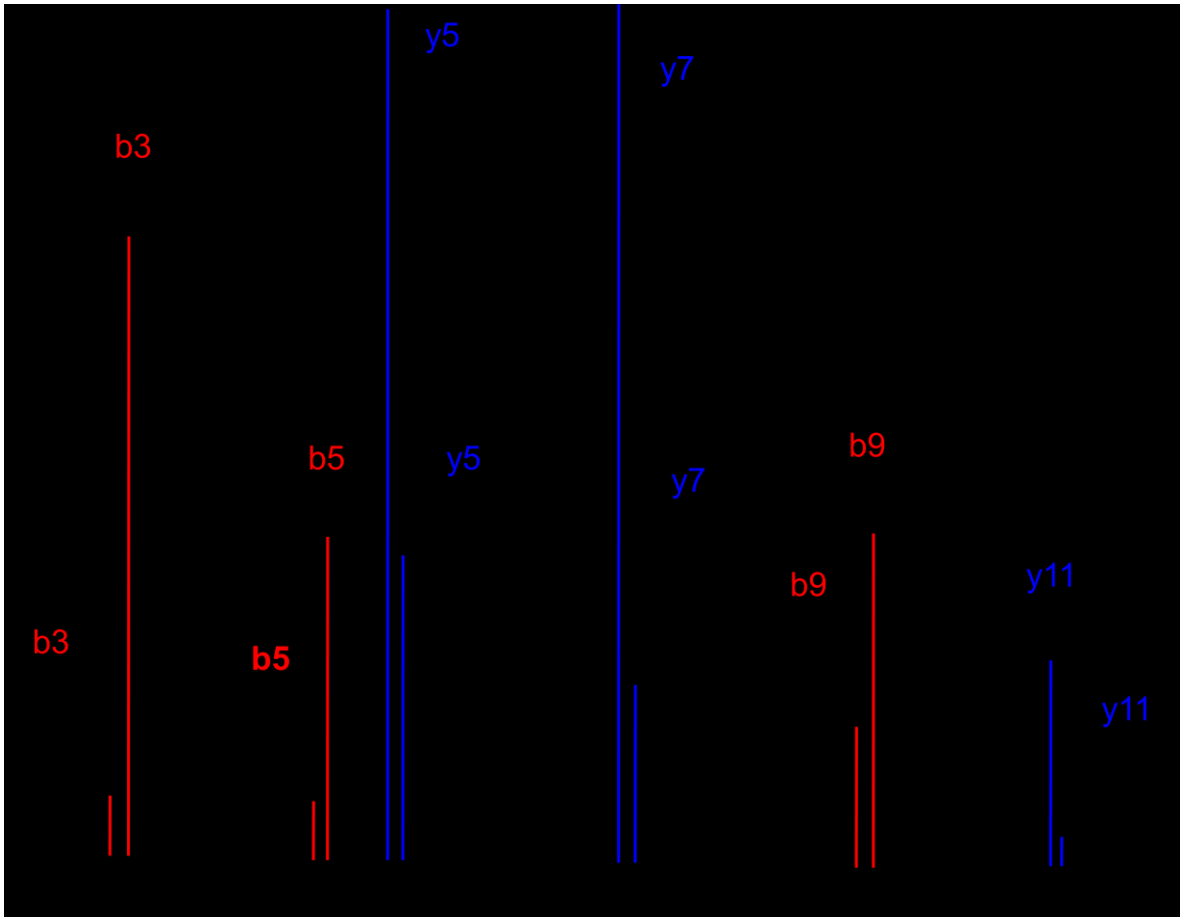


Figure 3-10. Expanded MS/MS spectral regions of monoacetylated GKGGKGLGKGGAKR. The mass spectral regions of the b-series (b_3, b_5 and b_9) and y-series (y_5, y_7 and y_{11}) ions used to quantitate the positional isomers of the monoacetylated (d_9) GKGGKGLGKGGAKR peptide are shown.

Table 3-2. Fractional abundances of fragment ions of mono-acetylated GKGGKGLGKGGAKR peptide from histone H4 from HeLa cells.				
Fragment ions	acK5	acK8	acK12	acK16
y5 and b9	0.256			0.744
y7 and b5	0.127		0.873	
y11 and b3	0.073	0.927		
Untreated HeLa cells	7%	5%	13%	74%
y5 and b9	0.303			0.697
y7 and b5	0.146		0.854	
y11 and b3	0.063	0.937		
TSA-treated HeLa cells	6%	8%	16%	70%

There are six potential isomeric forms of the diacetylated H4 tail peptide$_{4-17}$, which can be deciphered by computing relative abundances from the fragment ions of the MS1 isoform peak. Although it is statistically not possible to distinguish all the diacetylated isomeric species from unique fragment masses, we determined that about 6% of untreated diacetylated peptide$_{4-17}$ was acetylated at the K5 and K8 residues, whereas 20% were mixtures of acK_5acK_{12} and acK_5acK_{16}, 34% comprised mixtures of acK_8acK_{12} and acK_8acK_{16}, and 40% were acK_{12}acK_{16} (Table 3-3). When the deacetylase inhibitor-treated histone H4 sample was evaluated, the most dominant diacetylated isomer was the acK_{12}acK_{16} peptide, which constituted 59% of all diacetylated species. Three percent of the treated, diacetylated peptide$_{4-17}$ was acK_5acK_8, 14% was the combination of acK_5acK_{12} and acK_5acK_{16}, and 24% was the combination of acK_8acK_{12} and acK_8acK_{16}. Results for the tri-acetylated peptide species are summarized in Table 3-4. In untreated samples, 23% of the triacetylated peptide$_{4-17}$ was acK_5acK_8acK_{12}, 10% was acK_5acK_8acK_{16}, 28% was acK_5acK_{12}acK_{16}, and acK_8acK_{12}acK_{16} comprised 39%. For the trichostatin A/nitotinamide treated samples, 6% of all triacetylated peptide$_{4-17}$ species were acK_5acK_8 acK_{12}, 11% of the total was acK_5acK_8acK_{16}, 20% was acK_5acK_{12}acK_{16} and 62% was acK_8acK_{12}acK_{16}.

Table 3-3. Fractional abundances of fragment ions of **di-acetylated** GKGGKGLGKGGAKR peptide from histone H4 from HeLa cells.						
Fragment ions	acK5 acK8	acK5 acK12	acK5 acK16	acK8 acK12	acK8 acK16	acK12 acK16
y7 and b5	0.056	0.541				0.402
y11 and b3	0.255			0.745		
Untreated HeLa cells	6%	20%		34%		40%
y7 and b5	0.032	0.373				0.595
y11 and b3	0.168			0.832		
TSA-treated HeLa cells	3%	14%		24%		59%

Table 3-4. Fractional abundances of fragment ions of **tri-acetylated** GKGGKGLGKGGAKR peptide from histone H4 from HeLa cells.				
Fragment ions	acK5 acK8 acK12	acK5 acK8 acK16	acK5 acK12 acK16	acK8 acK12 acK16
b9 and y5	0.231	0.768		
y7 and b5	0.327		0.673	
y11 and b3	0.607			0.392
Untreated HeLa cells	23%	10%	28%	39%
b9 and y5	0.064	0.936		
y7 and b5	0.174		0.826	
y11 and b3	0.379			0.621
TSA-treated HeLa cells	6%	11%	20%	62%

Combining the distributions of Tables 3-2, 3-3, and 3-4, we can determine the relative abundance for all positional isomers. Thus Table 3-5 displays the relative abundance of all 16 acetylated isoforms for the tail peptide. Without deacetylase inhibition, the most abundant peptide isoform is unacetylated (63%). The second most abundant isoform had a single acetylation at residue K16 (22%). In the presence of deacetylase inhibitors, as would be predicted, the most abundant peptide is fully acetylated (32%). Compared to untreated controls, no single modification profile predominates; however, isoforms that include modifications at K16 are generally more abundant.

Table 3-5. Relative abundance of all 16 positional differentially modified peptides from HeLa cells comparing presence of deacetylase inhibitors. HeLa cells treated with deacetylase inhibitors TSA and NIA were analyzed for site-specific modifications on the histone H4 peptide GKGGKGLGKGGAKR. Lysine residues in red indicate acetylation.

	Isoforms	No de-acetylase inhibitors	With de-acetylase inhibitors
0	G **K** G G **K** G L G **K** G G A **K** R	**63%**	**8%**
1 acetylation	**G K G G K G L G K G G A K R**	**2%**	**1%**
	G K **G G K G L G K G G A K R**	**1%**	**1%**
	G K G G K G L G K **G G A K R**	**4%**	**2%**
	G K G G K G L G K G G A K **R**	**22%**	**10%**
2 acetylations	**G K G G K G L G K G G A K R**	**0.4%**	**0.6%**
	G K G G K G L G K G G A K R	**1%**	**3%**
	G K G G K G L G K G G A K R		
	G K G G K G L G K G G A K R	**2%**	**5%**
	G K G G K G L G K G G A K R		
	G K G G K G L G K G G A K R	**3%**	**12%**
3 acetylations	**G K G G K G L G K G G A K R**	**0.20%**	**2%**
	G K G G K G L G K G G A K R	**0.10%**	**3%**
	G K G G K G L G K G G A K R	**0.30%**	**5%**
	G K G G K G L G K G G A K R	**0.40%**	**16%**
4	**G K G G K G L G K G G A K R**	**0.30%**	**32%**

Methylation of the K_{79} site in histone H3. It is more difficult to design an analogous strategy that uses isotopically labeled methylation to elucidate methylation sites for two reasons: because lysines may be mono-, di- and tri-methylated and because methylation can also occur on arginine residues. Figure 3-11 shows MALDI TOF mass spectra of the yeast histone H3 fragment EIAQDFKTDLR corresponding to residues 73-83. An incomplete tryptic digestion of underivatized wild-type histone H3 is shown (Figure. 3-11a), revealing peaks corresponding to peptides unmodified, methylated, dimethylated and trimethylated (or acetylated which is found at the same *m/z*) at K_{79}. Tryptic digestion of the deuteroacetylated histone H3 fragment in Figure 3-11b shows both the unmethylated and mono-methylated species have been derivatized, while the dimethylated and trimethylated (or acetylated) species are not derivatized. Specifically the unmodified peptide seen at *m/z* 1335.85 in Figure 3-11a now appears at *m/z* 1380.97 as the deuteroacetylated species three mass units above the trimethylated (or acetylated) species at *m/z* 1377.82. Derivatization of the mono-methyl species is observed in Figure 3-11b at *m/z* 1386.99.

Figure 3-11. **MALDI TOF mass spectrum of the methylated yeast histone H3 73-83 peptide EIAQDFKTDLR**. The spectrum shows (a) unmethylated peptides and peptides methylated, dimethylated and trimethylated at lysine 79 and (b) deuteroacetylated peptide showing derivatization of both unmethylated and monomethylated species.

Electrospray mass spectra were also obtained for a yeast histone H3 sample containing unmodified and methylated K_{79} in part to differentiate trimethylated and acetylated species by high mass resolution of the Orbitrap. MS/MS spectra are shown in Figure 3-12, where the high mass accuracy obtainable from the Orbitrap mass spectrometer is noted for several of the major fragment ions. A deuteroacetylated tryptic digestion contained methylated/deuteroacetylated (a) and deuteroacetylated (b) species. Interestingly, modification of the basic lysine residue in both cases produces fragmentation patterns that are quite equivalent. On the chromatographic time frame (Figure 3-13), the retention times are similar, though not identical. Integration of the mass spectra across this range of retention time produces the composite spectrum shown in Figure 3-14.

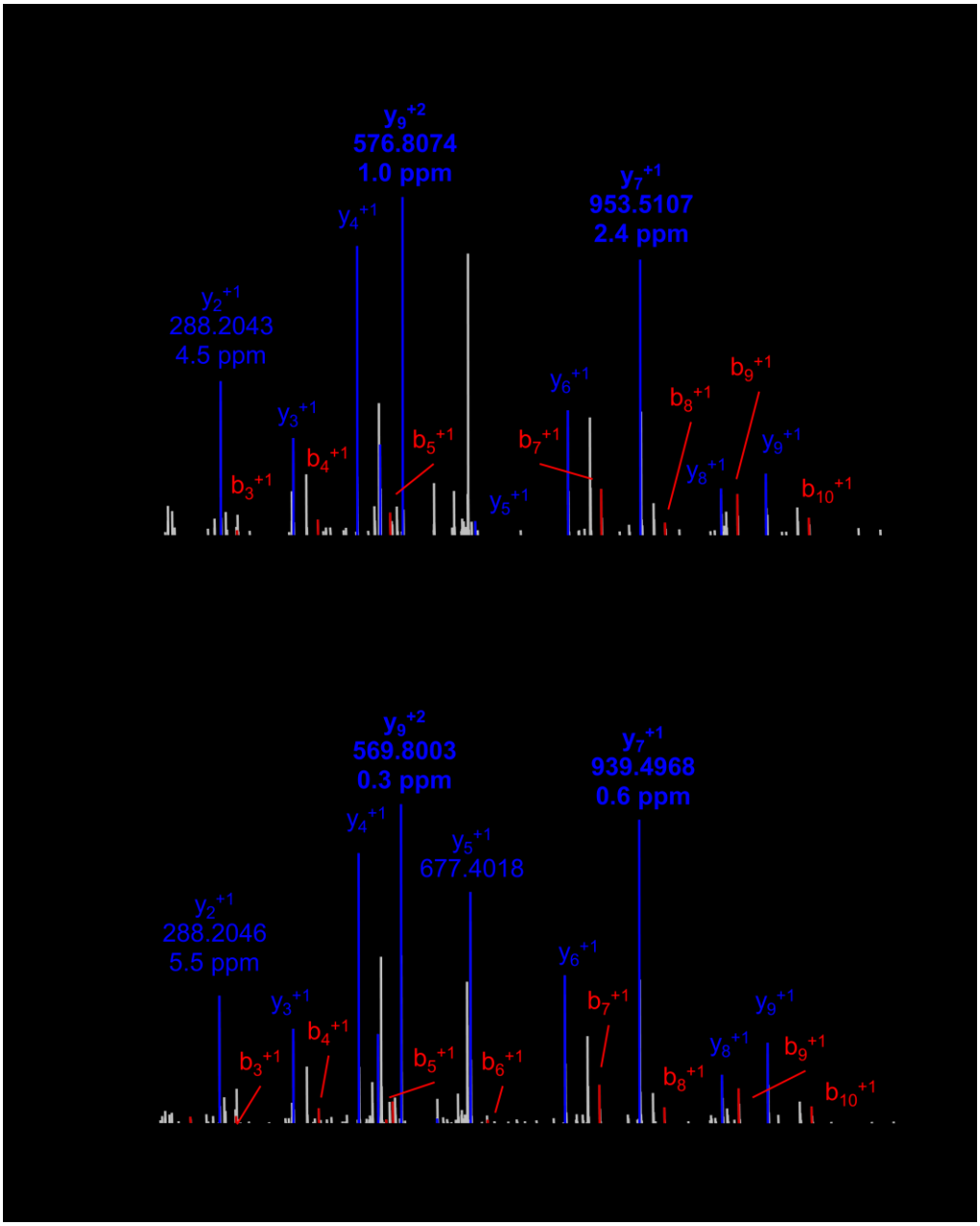


Figure 3-12. ESI Orbitrap MS/MS spectra of methylated EIAQDFKTDLR. The doubly-charged molecular ions of the derivatized peptide EIAQDFKTDLR is shown to be (a) methylated at lysine 79 and (b) unmethylated. The different masses of the y_7 and y_9 ions distinguish the two species. Mass accuracy is shown for the major fragment peaks.

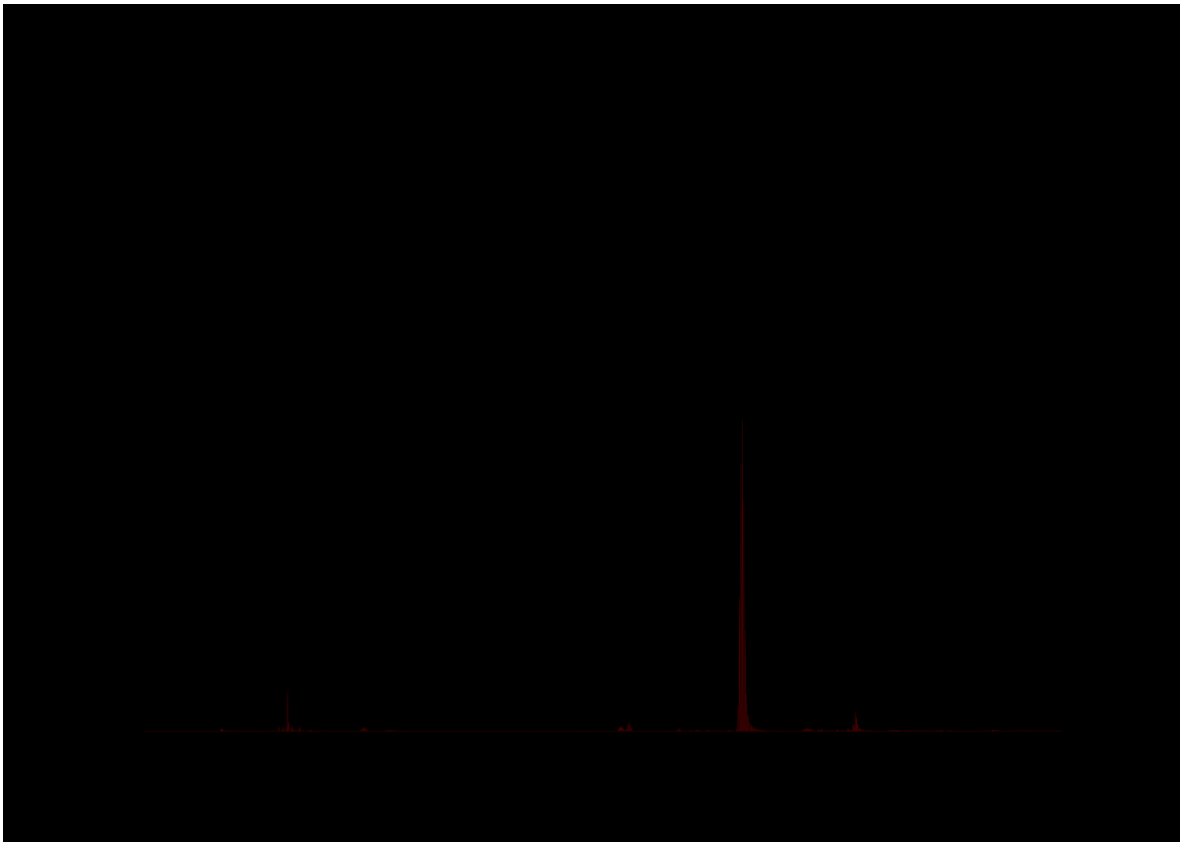


Figure 3-13. Methylated and unmethylated histone H3 EIAQDFKTDLR peptide HPLC chromatographs. C18 HPLC single ion chromatographs of derivatized (a) unmethylated and (b) methylated EIAQDFKTDLR obtained by monitoring the doubly charged molecular ions.

Figure 3-14. Methylated and unmethylated histone H3 EIAQDFKTDLR spectral integration for quantitation. Elution peaks of methylated and unmethylated EIAQDFKTDLR peptide were integrated to generate spectrum obtained across the retention times of the two species.

Methylation of histone H3 peptide KSAPSTGGVKKPHR. A tryptic fragment at *m/z* 533.6494 corresponds to the expected mass of the triply-charged peptide ion from the yeast $H3_{27\text{-}40}$ peptide KSAPSTGGVKKPHR carrying a methyl group and three deuteroacetylated lysines. The arginine residue and all three lysines are potential methylation sites. The MS/MS spectrum of the triply-charged molecular ion is shown in Figure 3-15. All of the observed b-series and y-series ions support methylation at either the K_{36} or K_{37} residues. The fragment ion that should distinguish these two possible structures is the singly-charged y_4 ion, where the major peak at *m/z* 582.3545 (Figure 3-15 inset) corresponds to methylation of the K_{36} lysine. The mass accuracy of 0.8 ppm obtainable on the Orbitrap mass analyzer is excellent and consistent with the mass accuracy in that range for other observed fragment ions. A smaller peak at *m/z* 596.3589 may correspond to the y_4 ion of a structure in which the K_{37} site is methylated, although the low intensity, signal to noise ratio, and low mass accuracy at 20 ppm make this peak's identity inconclusive and insignificant.

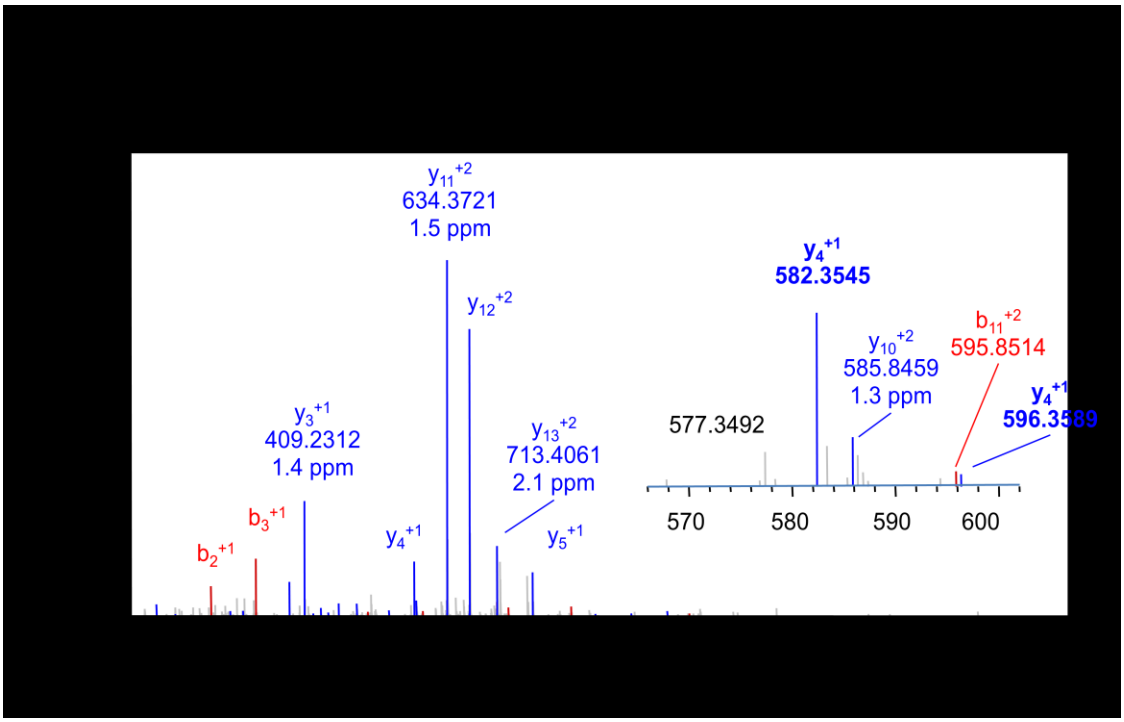


Figure 3-15. ESI Orbitrap mass spectrum of the molecular ion of the singly methylated and fully deuteroacetylated H3 peptide KSAPSTGGVKKPHR. Both the fragmentation pattern and mass accuracy are used to determine the location of the methylated lysine. The major y_4 fragment ion at 582.3545 differs 0.8 ppm from the calculated mass 582.3550 of isoform having methylation at the second lysine. A smaller peak at 596.3589 is close to the expected mass for the y_4 ion of the isoform having methylation at the third lysine, but has an error of 20 ppm.

Trimethylation of histone H3 peptide KSAPSTGGVKKPHR. The deuteroacetylated histone $H3_{27\text{-}40}$ peptide from HeLa cells also reveals a trimethylation site. The MS/MS spectrum of the triply-charged molecular ion observed at *m/z* 527.9819 is shown in Figure 3-16, and shows the same characteristic ions corresponding to b_2^{+1}, b_3^{+1} and y_{11} to y_{13}. These fragment masses are consistent with a structure having two deuteroacetyled lysine residues and a trimethylated lysine residue, and are accurate within 2-3 ppm. In contrast, errors in the mass accuracies when compared with a structure having two deuteroacetylated and one acetylated lysine are of the order of 20-30 ppm, the approximate mass difference between an acetyl and a trimethyl modification. The masses of the $b_2^{+1}$and b_3^{+1} ions preclude trimethylation on the K_{27} residue, but no fragment ion was observed that would distinguish between trimethylation at the K_{36} and K_{37} sites. Based upon the observation of methylation at K_{36} in the prior example then, the likely structure is KSAPSTGGVK_{3Me}KPHR (observed as K_{dAc}SAPSTGGV$K_{3Me}K_{dAc}$PHR).

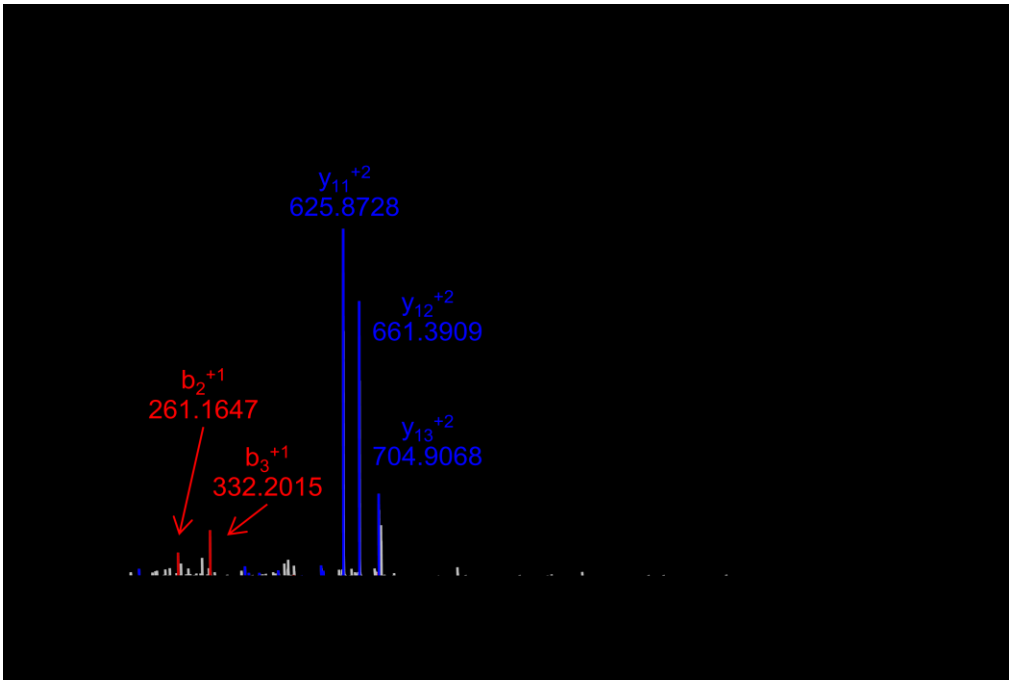


Figure 3-16. High resolution ESI Orbitrap MS/MS spectrum of KSAPSTGGVKKPHR peptide distinguishes acetylation from trimethylation. The inset table lists the high intensity ions observed for PTM identification, the calculated mass from the peptide listed above, and the difference in mass between calculated and observed. Differentiation between acetyl and trimethyl masses requires less than 30 PPM accuracy.

Discussion

There is considerable interest in developing global mass spectrometry approaches to assess lysine acetylation, particularly as this type of modification now appears to play significant roles in many cellular processes and species. We demonstrate our quantification method using bottom up derivatization for acetyl-lysine quantification and non-acetylation identification on histones using high resolution mass spectrometry. There are sixteen potential positional isomeric forms of the acetylated/deuteroacetylated peptide $GK_5GGK_8GLGK_{12}GGAK_{16}R$, including four mono-, six di- and four tri-acetylated species. Quantitation of positional isomers having the same precursor mass could not be determined using MALDI tandem analysis. Time-of-flight instruments generally have somewhat limited mass selection capability, so one cannot select isotopically pure precursors. Nonetheless, our preliminary work with this instrument indicated that the major monoacetylated species was acetylated at K16, while the major diacetylated species was acetylated on K8 and K16 (Cotter et al, 2007). This of course motivated our use of the LTQ/Orbitrap mass spectrometer to provide detailed analysis of the positional isomers.

The yeast histones analyzed in this study confirm previous results from this lab that $K_{16}Ac$ is the major mono-acetylated isoform (Cotter et al, 2007). Our findings are also consistent with previous biological studies. Histone H4 K16 acetylation is a reversible modification implicated in the widespread process of chromatin condensation by recruiting HATs to acetylate K12, K8, and K5. In mammals, K16

acetylation also correlates with K20 trimethylation (Sarg et al, 2004). Using an antibody approach to study site specific lysine acetylation in yeast histone H4, Suka et al. (Suka et al, 2001) found that lysine acetylation in the H4 N-terminal peptide$_{4\text{-}17}$ occurs sequentially from the C-terminus to the N-terminus: K16 acetylation precedes K12Ac, which precedes K8Ac, leading finally to K5Ac. We found similar results for H4 tail peptides from TSA/NIA treated HeLa cells: 6% of the monoacetylated peptide was acetylated at the K5 residue, 8% at the K8 residue, 16% at K12, and 70% at K16 (Table 2).

Identification of the methylated peptide EIAQDFKTDLR in the deuteroacetylated sample was made simpler by the predictable trypsin cleavage pattern and the known mass shift. We identified that modification of either lysine residue with a methyl group in both cases produces equivalent fragmentation patterns, suggesting that their comparable ionization and fragmentation behavior may provide an opportunity to make some assessment of the degree of methylation. Absolute quantification is not possible due to the retention time and ionization differences between species, although our results suggest a possility for calculation of relative abundance.

Other groups such as Zhao et al. (Zhang et al, 2009) have used alternative methods to quantitate lysine acetylation. Their approach utilizes anti-acetyl lysine antibodies to enrich for endogenously acetylated proteins and stable isotope labeling with amino acids in cell culture (SILAC) for quantitation by mass spectrometry. The sirtuin proteins have been shown to have deacetylase activity

in mitochondria (Schwer et al, 2006). Another group, Kelleher et al. (Li et al, 2009), report a method for global histone profiling in response to inhibition or knockdown of human deacetylases using a linear ion trap fourier transform mass spectrometer. In this approach, histone mixtures are subjected to methionine oxidation prior to chromatographic separation to enable resolution of each histone type (H1, H2.B, H2A-1, etc.). The high mass resolution and accuracy of the fourier transform mass spectrometer then enable one to determine the numbers of acetyl and methyl groups, and the method generally regards these to be the most abundant isoforms, e.g. H4+2Me occurs on H4K20, H4+2Me +2Ac occurs as $H4K20_{2Me}K12_{Ac}K16_{Ac}$, etc. This approach has the advantage of assessing all of the possible modification types, provides some quantitation of the overall modifications, but does not use MS/MS to provide details of positional isomeric forms. Alternatively, top down methods using electron capture dissociation on the fourier transform instrument (Siuti & Kelleher, 2007) provide structural verification of the major isoforms, but are not necessarily as quantitative as isotope based methods.

Our interest in the bottom up approach and chemical derivatization is based upon the possibility for exquisite and facile quantification of acetylation at specific residues, which was established in an earlier study by us of the acetylation of the K_{56} site in histone H3 for a series of *hst3* and *hst4* mutants (Celic et al, 2006). In that study, MALDI mass spectra of protein digests (as shown in Figure 9) were obtained for histones derived from wild type, deletion mutants and H184A,

N152A and D154A mutants. Using this quantitative method for cells synchronized in G1 phase, we were then able to monitor changes in K_{56} acetylation to observe the inactivation of Hst3/Hst4p during passage through S Phase in response to NIA. Regulation of histone H3 K_{56} has been shown to be critical in the cell cycle of fungi (Xhemalce et al, 2007) and is regulated by the fungal histone acetyltransferase Rtt109 (Tang et al, 2008), a homolog of p300/CBP specific for K_{56}. In addition to the peptide containing K_{56}, expansion of the isotopic displays from the MALDI mass spectra enabled quantitation of the acetylation of lysine residues in K_9STGGK_{14}APR and K_{18}QLASK_{23}AAR in histone H3.

In addition to obtaining quantitative information on acetylation at the positional isoform level, the intent is to begin to utilize this bottom up derivatization method to provide qualitative and semi-quantitative analyses for other modifications, particularly methylation, dimethylation and trimethylation. Methylation of the K_{79} site on histone H3 is of particular importance in cell cycle and replication, and Kelleher et al. (Sweet et al, 2010) have recently used a stable isotope strategy (SILAC) and mass spectrometry to compare the methylation and dimethylation of pre-existing and newly-synthesized histones. While it is clear that the deuteroacetylated structures: EIAQDFK_{Me+dAc}TDLR and EIAQDFK_{dAc}TDLR encompassing the K_{79} site are not chemically and chromatographical identical, as in the case of stable isotope analogs, their similar responses to ionization and mass spectral sensitivity suggest that with the proper peptide analog experiment,

analyses intended to observe changes in methylation at this site could be easily calibrated for accurate quantitation.

A bottom up approach using global chemical derivatization of endogenously unmodified lysines is an effective alternative to other mass spectrometric histone analysis methods in quantification of site-specific histone acetylation. Finally, this method could be extended to facile identification of modifications other than acetylation.

Chapter 4:

Conclusions

The field of proteomics is expanding rapidly as new proteomic techniques are being developed and novel proteomic applications are expanding biological research. The work presented in the first part of the thesis demonstrates the application of the established SRM technique to describe natural variation of a highly similar protein family in a novel population where other methods of protein quantification were impractical. This research will serve as a resource for many researchers interested in the murine Cyp pharmacology. Furthermore, introduction of proteomics as a biological tool for evaluation of global Cyp expression differences in the field of Cyp biology expands the field of proteomics by the exposure of basic biologists to the powerful proteomic techniques available. The Cyp SRM assay presented in this research could be applied to determine the dynamics of Cyp expression in a disease model or induction experiment.

Because few studies aim to make functional correlations with human proteins and limited information is available regarding when and where murine proteins are expressed, the Cyp expression results presented in this thesis are exclusively descriptive. Murine Cyp activity could be correlated to Cyp protein expression by activity assays with pure Cyp isozymes. This is done using recombinant bacteria to purify a single Cyp and incubating each Cyp with substrates known to be nearly exclusively metabolized by a particular Cyp and other substrates of interest. Establishing murine Cyps functionally homologous

to well-characterized Cyps through metabolism of a common substrate will make the proteomics work presented here more accessible to researchers in all basic science fields.

Because the cytochrome P450 family is highly homologous, the choice of peptides for protein quantitation was limited for certain isozymes, resulting in variability in peak intensity. Further Cyp proteomics experiments would benefit from improving results statistics. This could be done by establishing individual variability by analyzing protein from individuals instead of pooled sample. Further, reproducibility of peptides with lower ionization efficiency could be enhanced with replicate injections. All peptides do not ionize equally efficiently, and for those that ionize less well, replicates display more variability. Three injection replicates in addition to the three technical replicates performed would drastically increase the amount of instrument time and therefore cost of the experiment, although they would significantly enhance the statistical power of the resulting data.

The Cyp SRM assay could be applied to humanized or knockout mouse models to test the induction of cytochrome P450s with elongated drug usage or in determining differences in Cyp expression. For example, Johnson et al. are investigating the role of 4A11 in hypertension using a humanized mouse model (Savas et al, 2009). In addition to characterizing changes in signaling pathways, they asked us to employ the cytochrome P450 SRM assay to evaluate if

cytochrome P450 expression is also altered, and therefore, potentially implicated in a similar expression pathway as their CYP of interest.

By continuing to establish more accurate, quicker and easier quantitative methods, better tools for the field of mass spectrometry will be available for application to new biological questions. The PTM identification and quantification research discussed in Chapter 3 could be further established by validating quantification on methylated and acetylated synthetic peptides. Performing analysis on known mixtures of methylated and acetylated peptides would determine the precision of the quantitation calculation. In addition, one challenge in the application of the technique using LC-MS was to reduce sample complexity and background influence, as the high abundance of unacetylated peptides seemed to dominate the data-dependent MS/MS targeting. Analyzing the synthetic modified peptide mix in an E. coli digest would establish the robustness of the technique in a large mixture in comparison to traditional data-dependent searches. Furthermore, newer models of Orbitrap mass spectrometers have at least double the scanning speed, such that more peptides of interest might be targeted for fragmentation. These experiments could establish the technique as a validated system using any instrument or platform and any protein mixture.

The LTQ Orbitrap XL used for the experiments in Chapter 3 was one of the first Orbitrap models in academia, acquired in 2008. In the last 5 years, Orbitraps have become faster and more specialized, including the Velos with a dual-pressure linear ion trap to more than double the dynamic range at the same

scanning rate, the Elite with a high-field Orbitrap" which quadruples the resolution, the Exactive with a streamlined Orbitrap-only detection, and most recently the Fusion with a high field Orbitrap and two other detectors for the most flexibility in experimental design. A streamlined instrument for FT-FT detection with higher scan rate and higher resolution would be the instrument of choice for PTM quantification described here: the Elite, although both the histone PTM project and the cytochrome P450 project could have been conducted on the Thermo Fusion.

PTM identification work presented here did not include a practical way to identify peptides of unknown sequence that could potentially be modified. One method which I started to pursue at the beginning of my PTM identification work was to identify all acetylated peptides in comparison to the deuteroacetylated version of that peptide. It is standard to export a list of peak *m/z* and peak intensities, so I wrote an Excel Macro to eliminate masses with peak intensity below signal to noise, identify all peaks with a user-defined mass difference, and report all matches and peak intensities in a new excel sheet. This analysis was versatile, as the input mass difference could be either 3, 15, 12 or 27 for acetylation, methylation, dimethylation, or trimethylation, respectively. Although the macro identified all shifted peaks, the isotope elimination functions were ineffective at reducing false-positives, which made this analysis impractical. Another method to manually identify peaks corresponding to deuteroacetylated peptides in an unknown protein sample would be to compare mass spectra of a

deuteroacetylated sample to an identical but not deuteroacetylated sample would reveal one simple and one complex spectrum. Peaks that were more intense in the deuterated sample would indicate a lysine-containing peptides. After a peak has been identified as a K-containing peptide, all peaks of potential modifications including the mass shifts listed above can easily be targeted for MS/MS identification. The use of the MALDI-TOF instrument for this analysis instead of Orbitrap has the advantage of using approximately $1/20^{th}$ that of an Orbitrap analysis. In addition, all the information about the sample is contained in one spectrum, allowing for facile quantitation or relative abundance. Although these are interesting directions, these method optimizations have utility for a narrow subset of biological samples.

One research area that would benefit the whole field of proteomics would be to characterize the chemistry of peptide ionization. The chemical composition of the peptide and the whole protein is not an indicator for its success in mass spectrometric analysis, as it is not yet understood what chemistry composition in an analyte is best detected in a mass spectrometer (Mirzaei & Regnier, 2006). Currently, all peptides in a protein are experimentally evaluated for ionization efficiency. In cases such as the Cyp work presented here, the biological sample is also the sample used for method development. In the case of low abundant Cyps such as Cyp2u1, detection was noisy and peptide choice was not supported by any scientific findings regarding if it was predicted to be an adequate candidate for SRM. SRM analysis usually requires an expensive

synthetic peptide, so selection of a poor SRM candidate peptide can be a costly mistake. A prediction tool for SRM peptide selection will be a cost and time-saving tool for proteomics mass spectrometrists using triple quadrupole instruments.

The field of proteomics mass spectrometry is beginning to include biologists in addition to biology-inclined mass spectrometrists and chemists. The hundreds of megabytes of data generated per LC-MS analysis are available online by journals such as in Nature Methods and by academic groups such as PeptideAtlas. This exponentially growing field currently defines global protein identification and quantification, and the network of information it produces will provide a rich resource of global protein expression for mass spectrometrists and non-mass spectrometrists alike.

References

Abdelmegeed MA, Banerjee A, Yoo SH, Jang S, Gonzalez FJ, Song BJ (2012) Critical role of cytochrome P450 2E1 (CYP2E1) in the development of high fat-induced non-alcoholic steatohepatitis. *J Hepatol* **57:** 860-866

Ahmad Y, Lamond AI (2013) A perspective on proteomics in cell biology. *Trends Cell Biol*

Amunom I, Dieter LJ, Tamasi V, Cai J, Conklin DJ, Srivastava S, Martin MV, Guengerich FP, Prough RA (2011) Cytochromes P450 catalyze the reduction of alpha,beta-unsaturated aldehydes. *Chem Res Toxicol* **24:** 1223-1230

Avery LB, Parsons TL, Meyers DJ, Hubbard WC (2010) A highly sensitive ultra performance liquid chromatography-tandem mass spectrometric (UPLC-MS/MS) technique for quantitation of protein free and bound efavirenz (EFV) in human seminal and blood plasma. *J Chromatogr B Analyt Technol Biomed Life Sci* **878:** 3217-3224

Avery LB, Sacktor N, McArthur JC, Hendrix CW (2013) Protein-free efavirenz concentrations in cerebrospinal fluid and blood plasma are equivalent: applying the law of mass action to predict protein-free drug concentration. *Antimicrob Agents Chemother* **57:** 4095

Barlesi F, Giaccone G, Gallegos-Ruiz MI, Loundou A, Span SW, Lefesvre P, Kruyt FA, Rodriguez JA (2007) Global histone modifications predict prognosis of resected non small-cell lung cancer. *J Clin Oncol* **25:** 4358-4364

Bieche I, Narjoz C, Asselah T, Vacher S, Marcellin P, Lidereau R, Beaune P, de Waziers I (2007) Reverse transcriptase-PCR quantification of mRNA levels from cytochrome (CYP)1, CYP2 and CYP3 families in 22 different human tissues. *Pharmacogenet Genomics* **17:** 731-742

Bonisch C, Nieratschker SM, Orfanos NK, Hake SB (2008) Chromatin proteomics and epigenetic regulatory circuits. *Expert Rev Proteomics* **5:** 105-119

Boyne MT, 2nd, Pesavento JJ, Mizzen CA, Kelleher NL (2006) Precise characterization of human histones in the H2A gene family by top down mass spectrometry. *J Proteome Res* **5:** 248-253

Celic I, Masumoto H, Griffith WP, Meluh P, Cotter RJ, Boeke JD, Verreault A (2006) The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Curr Biol* **16:** 1280-1289

Center for Drug Evaluation and Research (CDER) USFaDAF. (2001) Guidance for Industry: bioanalytical method validation.

Cheung P, Allis CD, Sassone-Corsi P (2000) Signaling to chromatin through histone modifications. *Cell* **103:** 263-271

Choudhary D, Jansson I, Schenkman JB, Sarfarazi M, Stoilov I (2003) Comparative expression profiling of 40 mouse cytochrome P450 genes in embryonic and adult tissues. *Arch Biochem Biophys* **414:** 91-100

Choudhary D, Jansson I, Stoilov I, Sarfarazi M, Schenkman JB (2005) Expression patterns of mouse and human CYP orthologs (families 1-4) during development and in different adult tissues. *Arch Biochem Biophys* **436:** 50-61

Clancy T, Hovig E (2013) From proteomes to complexomes in the era of systems biology. *Proteomics*

Cohen SL, Chait BT (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal Chem* **68:** 31-37

Cornish TJ, Cotter RJ (1993) A curved-field reflectron for improved energy focusing of product ions in time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **7:** 1037-1040

Cotter RJ, American Chemical Society. Meeting (1994) *Time-of-flight mass spectrometry*, Washington, DC: American Chemical Society.

Cotter RJ, Gardner BD, Iltchenko S, English RD (2004) Tandem time-of-flight mass spectrometry with a curved field reflectron. *Anal Chem* **76:** 1976-1981

Cotter RJ, Griffith W, Jelinek C (2007) Tandem time-of-flight (TOF/TOF) mass spectrometry and the curved-field reflectron. *J Chromatogr B Analyt Technol Biomed Life Sci* **855:** 2-13

Cressey TR, Stek A, Capparelli E, Bowonwatanuwong C, Prommas S, Sirivatanapa P, Yuthavisuthi P, Neungton C, Huo Y, Smith E, Best BM, Mirochnick M (2012) Efavirenz pharmacokinetics during the third trimester of pregnancy and postpartum. *J Acquir Immune Defic Syndr* **59:** 245-252

Dekant W, Frischmann C, Speerschneider P (1995) Sex, organ and species specific bioactivation of chloromethane by cytochrome P4502E1. *Xenobiotica* **25:** 1259-1265

Dhir RN, Dworakowski W, Thangavel C, Shapiro BH (2006) Sexually dimorphic regulation of hepatic isoforms of human cytochrome p450 by growth hormone. *J Pharmacol Exp Ther* **316:** 87-94

Ding X, Kaminsky LS (2003) Human extrahepatic cytochromes P450: function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. *Annu Rev Pharmacol Toxicol* **43:** 149-173

Flint MS, Hood BL, Sun M, Stewart NA, Jones-Laughner J, Conrads TP (2010) Proteomic analysis of the murine liver in response to a combined exposure to psychological stress and 7,12-dimethylbenz(a)anthracene. *J Proteome Res* **9:** 509-520

Garcia BA, Barber CM, Hake SB, Ptak C, Turner FB, Busby SA, Shabanowitz J, Moran RG, Allis CD, Hunt DF (2005) Modifications of human histone H3 variants during mitosis. *Biochemistry* **44:** 13202-13213

Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, Allis CD, Hunt DF (2007a) Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* **282:** 7641-7655

Garcia BA, Mollah S, Ueberheide BM, Busby SA, Muratore TL, Shabanowitz J, Hunt DF (2007b) Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat Protoc* **2:** 933-938

Garcia BA, Pesavento JJ, Mizzen CA, Kelleher NL (2007c) Pervasive combinatorial modification of histone H3 in human cells. *Nat Methods* **4:** 487-489

Garcia BA, Shabanowitz J, Hunt DF (2007d) Characterization of histones and their post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* **11:** 66-73

Hernandez JP, Chapman LM, Kretschmer XC, Baldwin WS (2006) Gender-specific induction of cytochrome P450s in nonylphenol-treated FVB/NJ mice. *Toxicol Appl Pharmacol* **216:** 186-196

Hopfgartner G, Varesio E, Tschappat V, Grivet C, Bourgogne E, Leuthold LA (2004) Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *J Mass Spectrom* **39:** 845-855

Jager M, Deechongkit S, Koepf EK, Nguyen H, Gao J, Powers ET, Gruebele M, Kelly JW (2008) Understanding the mechanism of beta-sheet folding from a chemical and biological perspective. *Biopolymers* **90:** 751-758

Jenkins RE, Kitteringham NR, Hunter CL, Webb S, Hunt TJ, Elsby R, Watson RB, Williams D, Pennington SR, Park BK (2006) Relative and absolute quantitative expression profiling of cytochromes P450 using isotope-coded affinity tags. *Proteomics* **6:** 1934-1947

Jiang L, Smith JN, Anderson SL, Ma P, Mizzen CA, Kelleher NL (2007) Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry. LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. *J Biol Chem* **282:** 27923-27934

Jimenez CR, Huang L, Qiu Y, Burlingame AL (2001) In-gel digestion of proteins for MALDI-MS fingerprint mapping. *Curr Protoc Protein Sci* **Chapter 16:** Unit 16 14

Kalsotra A, Strobel HW (2006) Cytochrome P450 4F subfamily: at the crossroads of eicosanoid and drug metabolism. *Pharmacol Ther* **112:** 589-611

Koh KH, Xie H, Yu AM, Jeong H (2011) Altered cytochrome P450 expression in mice during pregnancy. *Drug Metab Dispos* **39:** 165-169

Kuska B (1998) Beer, Bethesda, and biology: how "genomics" came into being. *J Natl Cancer Inst* **90:** 93

Lawrence JH, Tuttle LW, Scott KG, Connor CL (1940) Studies on Neoplasms with the Aid of Radioactive Phosphorus. I. The Total Phosphorus Metabolism of Normal and Leukemic Mice. *J Clin Invest* **19:** 267-271

Lee DC, Stenland CJ, Hartwell RC, Ford EK, Cai K, Miller JL, Gilligan KJ, Rubenstein R, Fournel M, Petteway SR, Jr. (2000) Monitoring plasma processing steps with a sensitive Western blot assay for the detection of the prion protein. *J Virol Methods* **84:** 77-89

Li M, Jiang L, Kelleher NL (2009) Global histone profiling by LC-FTMS after inhibition and knockdown of deacetylases in human cells. *J Chromatogr B Analyt Technol Biomed Life Sci* **877:** 3885-3892

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251-260

Luo G, Zeldin DC, Blaisdell JA, Hodgson E, Goldstein JA (1998) Cloning and expression of murine CYP2Cs and their ability to metabolize arachidonic acid. *Arch Biochem Biophys* **357:** 45-57

MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26:** 966-968

Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* **72:** 1156-1162

Michaud GA, Salcius M, Zhou F, Bangham R, Bonin J, Guo H, Snyder M, Predki PF, Schweitzer BI (2003) Analyzing antibody specificity with whole proteome microarrays. *Nat Biotechnol* **21:** 1509-1512

Mirzaei H, Regnier F (2006) Enhancing electrospray ionization efficiency of peptides by derivatization. *Anal Chem* **78:** 4175-4183

Morgan K, French SW, Morgan TR (2002) Production of a cytochrome P450 2E1 transgenic mouse and initial evaluation of alcoholic liver damage. *Hepatology* **36:** 122-134

Muller DN, Schmidt C, Barbosa-Sicard E, Wellner M, Gross V, Hercule H, Markovic M, Honeck H, Luft FC, Schunck WH (2007) Mouse Cyp4a isoforms: enzymatic properties, gender- and strain-specific expression, and role in renal 20-hydroxyeicosatetraenoic acid formation. *Biochem J* **403:** 109-118

Nakagawa K, Holla VR, Wei Y, Wang WH, Gatica A, Wei S, Mei S, Miller CM, Cha DR, Price E, Jr., Zent R, Pozzi A, Breyer MD, Guan Y, Falck JR, Waterman MR, Capdevila JH (2006) Salt-sensitive hypertension is associated with dysfunctional Cyp4a10 gene and kidney epithelial sodium channel. *J Clin Invest* **116:** 1696-1702

Nebert DW, Nelson DR, Adesnik M, Coon MJ, Estabrook RW, Gonzalez FJ, Guengerich FP, Gunsalus IC, Johnson EF, Kemper B, et al. (1989) The P450 superfamily: updated listing of all genes and recommended nomenclature for the chromosomal loci. *DNA* **8:** 1-13

Nelson DR, Strobel HW (1987) Evolution of cytochrome P-450 proteins. *Mol Biol Evol* **4:** 572-593

Nelson DR, Zeldin DC, Hoffman SM, Maltais LJ, Wain HM, Nebert DW (2004) Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* **14:** 1-18

Olsen JV, Ong SE, Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **3:** 608-614

Parkinson OT, Liggitt HD, Rettie AE, Kelly EJ (2013) Generation and characterization of a Cyp4b1 null mouse and the role of CYP4B1 in the activation and toxicity of Ipomeanol. *Toxicol Sci* **134:** 243-250

Penaloza CG, Estevez B, Han DM, Norouzi M, Lockshin RA, Zakeri Z (2013) Sex-dependent regulation of cytochrome P450 family members Cyp1a1, Cyp2e1, and Cyp7b1 by methylation of DNA. *FASEB J*

Pesavento JJ, Kim YB, Taylor GK, Kelleher NL (2004) Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry. *J Am Chem Soc* **126:** 3386-3387

Pham MH, Rhinn H, Auzeil N, Regazzetti A, Harami DE, Scherman D, Chabot GG (2011) Identification and induction of cytochrome P450s involved in the metabolism of flavone-8-acetic acid in mice. *Drug Metab Lett* **5:** 73-84

Plazas-Mayorca MD, Zee BM, Young NL, Fingerman IM, LeRoy G, Briggs SD, Garcia BA (2009) One-pot shotgun quantitative mass spectrometry characterization of histones. *J Proteome Res* **8:** 5367-5374

Renaud HJ, Cui JY, Khan M, Klaassen CD (2011) Tissue distribution and gender-divergent expression of 78 cytochrome P450 mRNAs in mice. *Toxicol Sci* **124:** 261-277

Rodriguez-Antona C, Donato MT, Pareja E, Gomez-Lechon MJ, Castell JV (2001) Cytochrome P-450 mRNA expression in human liver and its relationship with enzyme activity. *Arch Biochem Biophys* **393:** 308-315

Sarg B, Helliger W, Talasz H, Koutzamani E, Lindner HH (2004) Histone H4 hyperacetylation precludes histone H4 lysine 20 trimethylation. *J Biol Chem* **279:** 53458-53464

Savas U, Machemer DE, Hsu MH, Gaynor P, Lasker JM, Tukey RH, Johnson EF (2009) Opposing roles of peroxisome proliferator-activated receptor alpha and growth hormone in the regulation of CYP4A11 expression in a transgenic mouse model. *J Biol Chem* **284:** 16541-16552

Schirmer M, Rosenberger A, Klein K, Kulle B, Toliat MR, Nurnberg P, Zanger UM, Wojnowski L (2007) Sex-dependent genetic markers of CYP3A4 expression and activity in human liver microsomes. *Pharmacogenomics* **8:** 443-453

Schwer B, Bunkenborg J, Verdin RO, Andersen JS, Verdin E (2006) Reversible lysine acetylation controls the activity of the mitochondrial enzyme acetyl-CoA synthetase 2. *Proc Natl Acad Sci U S A* **103:** 10224-10229

ShimadzuBiotech. (2006) AXIMA-TOF2: Specification.

Shinde SS, Hay MP, Patterson AV, Denny WA, Anderson RF (2009) Spin trapping of radicals other than the *OH radical upon reduction of the anticancer agent tirapazamine by cytochrome P450 reductase. *J Am Chem Soc* **131:** 14220-14221

Siuti N, Kelleher NL (2007) Decoding protein modifications using top-down mass spectrometry. *Nat Methods* **4:** 817-821

Siuti N, Roth MJ, Mizzen CA, Kelleher NL, Pesavento JJ (2006) Gene-specific characterization of human histone H2B by electron capture dissociation. *J Proteome Res* **5:** 233-239

Smith CM (2005) Quantification of acetylation at proximal lysine residues using isotopic labeling and tandem mass spectrometry. *Methods* **36:** 395-403

Smith CM, Gafken PR, Zhang Z, Gottschling DE, Smith JB, Smith DL (2003) Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal Biochem* **316:** 23-33

Stiborova M, Dracinska H, Martinek V, Svaskova D, Hodek P, Milichovsky J, Hejdukova Z, Brotanek J, Schmeiser HH, Frei E (2013) Induced expression of cytochrome P450 1A and NAD(P)H:quinone oxidoreductase determined at mRNA, protein, and enzyme activity levels in rats exposed to the carcinogenic azo dye 1-phenylazo-2-naphthol (Sudan I). *Chem Res Toxicol* **26:** 290-299

Su X, Ren C, Freitas MA (2007) Mass spectrometry-based strategies for characterization of histones and their post-translational modifications. *Expert Rev Proteomics* **4:** 211-225

Suka N, Suka Y, Carmen AA, Wu J, Grunstein M (2001) Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell* **8:** 473-479

Suter L, Babiss LE, Wheeldon EB (2004) Toxicogenomics in predictive toxicology in drug development. *Chem Biol* **11:** 161-171

Sutton CW, Sutherland M, Shnyder S, Patterson LH (2010) Improved preparation and detection of cytochrome P450 isoforms using MS methods. *Proteomics* **10:** 327-331

Sweet SM, Li M, Thomas PM, Durbin KR, Kelleher NL (2010) Kinetics of re-establishing H3K79 methylation marks in global human chromatin. *J Biol Chem* **285:** 32778-32786

Tang Y, Holbert MA, Wurtele H, Meeth K, Rocha W, Gharib M, Jiang E, Thibault P, Verreault A, Cole PA, Marmorstein R (2008) Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP. *Nat Struct Mol Biol* **15:** 998

Tateishi T, Nakura H, Asoh M, Watanabe M, Tanaka M, Kumai T, Takashima S, Imaoka S, Funae Y, Yabusaki Y, Kamataki T, Kobayashi S (1997) A comparison of hepatic cytochrome P450 protein expression between infancy and postinfancy. *Life Sci* **61:** 2567-2574

ThermoFisherScientific. (2009a) Product Specification: Thermo Scientific TSQ Vantage.

ThermoFisherScientific. (2009b) Product Specifications: Thermo Scientific LTQ Orbitrap Velos.

Tian R (2013) Exploring intercellular signaling by proteomic approaches. *Proteomics*

Tyers M, Mann M (2003) From genomics to proteomics. *Nature* **422:** 193-197

Vogel C, Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13:** 227-232

Wang AH, Bertos NR, Vezmar M, Pelletier N, Crosato M, Heng HH, Th'ng J, Han J, Yang XJ (1999) HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor. *Mol Cell Biol* **19:** 7816-7827

Wang H, Zhao Y, Bradbury JA, Graves JP, Foley J, Blaisdell JA, Goldstein JA, Zeldin DC (2004) Cloning, expression, and characterization of three new mouse cytochrome p450 enzymes and partial characterization of their fatty acid oxidation activities. *Mol Pharmacol* **65:** 1148-1158

Ward BA, Gorski JC, Jones DR, Hall SD, Flockhart DA, Desta Z (2003) The cytochrome P450 2B6 (CYP2B6) is the main catalyst of efavirenz primary and secondary metabolism: implication for HIV/AIDS therapy and utility of efavirenz as a substrate marker of CYP2B6 catalytic activity. *J Pharmacol Exp Ther* **306:** 287-300

Waterfield MD, Del Favero A, Gray CH (1969) Effect of 1,4-dihydro-3,5-dicarbethoxycollidine on hepatic microsomal haem, cytochrome b5 and cytochrome P450 in rabbits and mice. *Biochim Biophys Acta* **184:** 470-473

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520-562

Wienkers LC, Heath TG (2005) Predicting in vivo drug interactions from in vitro drug discovery data. *Nat Rev Drug Discov* **4:** 825-833

Wright AT, Cravatt BF (2007) Chemical proteomic probes for profiling cytochrome p450 activities and drug interactions in vivo. *Chem Biol* **14:** 1043-1051

Xhemalce B, Miller KM, Driscoll R, Masumoto H, Jackson SP, Kouzarides T, Verreault A, Arcangioli B (2007) Regulation of histone H3 lysine 56 acetylation in Schizosaccharomyces pombe. *J Biol Chem* **282:** 15040-15047

Xie X, Miao L, Yao J, Feng C, Li C, Gao M, Liu M, Gong L, Wang Y, Qi X, Ren J (2013) Role of multiple microRNAs in the sexually dimorphic expression of Cyp2b9 in mouse liver. *Drug Metab Dispos* **41:** 1732-1737

Xu M, Ju W, Hao H, Wang G, Li P (2013) Cytochrome P450 2J2: distribution, function, regulation, genetic polymorphisms and clinical significance. *Drug Metab Rev* **45:** 311-352

Young NL, Dimaggio PA, Garcia BA (2010) The significance, development and progress of high-throughput combinatorial histone code analysis. *Cell Mol Life Sci* **67:** 3983-4000

Young NL, DiMaggio PA, Plazas-Mayorca MD, Baliban RC, Floudas CA, Garcia BA (2009) High throughput characterization of combinatorial histone codes. *Mol Cell Proteomics* **8:** 2266-2284

Zhang J, Sprung R, Pei J, Tan X, Kim S, Zhu H, Liu CF, Grishin NV, Zhao Y (2009) Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli. *Mol Cell Proteomics* **8:** 215-225

Zhang QY, Dunbar D, Kaminsky LS (2003) Characterization of mouse small intestinal cytochrome P450 expression. *Drug Metab Dispos* **31:** 1346-1351

Curriculum Vitae

Elisabeth Marie Hersman

EDUCATION

2013 **PhD in Pharmacology and Molecular Sciences from John Hopkins School of Medicine** Baltimore, MD

2008 **B.A. in Chemistry and Neuroscience from Wellesley College** Wellesley, MA

2005- **Courses at Massachusetts Institute of Technology** Cambridge, MA

2007 Chemistry, Biochemistry and Pharmacology of Synaptic Transmission

RESEARCH EXPERIENCE

2013 **Thesis Research advisor Namandjé Bumpus**

2009 **Thesis Research advisor Robert Cotter** Baltimore, MD
Dr. Cotter passed away suddenly in November 2012

2008 **Research Rotation with advisor Richard Huganir**
Studied the interaction between the protein Tid1 and NMDA receptors to evaluate a novel role of Tid1 as a trafficking regulator

2008 **Research Rotation with advisor Solomon Snyder**
Knocked down PGDH protein in cortical glial cells to investigate its role in the D-serine synthesis pathway

2007 **Wellesley Summer Research Program advisor Nancy H. Kolodny**
Collected and analyzed in vivo and ex vivo NMR data studying the small molecule brain chemistry of a mouse model for Rett Sydrome

2006 **Brigham and Women's Hospital Summer Research advisor Samuel Patz** Boston, MA
Used hyperpolarized xenon to take MRIs of lungs, streamlining the technique in preparation for a COPD assessment study

GRANT SUPPORT

NIA Grant F31 AG041609-02 (PI: **E Hersman**) July 2012- July 2014
Quantifying Histone Modifications Associated with Age using Mass Spectrometry

PUBLICATIONS

Hersman E, Nelson DM, Griffith WP, Jelinek CA, Cotter RJ (2012)
Analysis of histone modifications from tryptic peptides of deuteroacetylated isoforms
International Journal of Mass Spectrometry. 312: **5–16.** PMID: 22389584

Manuscript submitted: **Hersman E,** Bumpus N (Dec 2013)
Profiling Murine Cytochrome P450 Expression Using a Targeted Proteomics Approach. Nature Biotechnology.

CONFERENCE PRESENTATIONS

2012 **Talk at World Human Proteome Organization Conference**
Studying Age and Calorie Restriction Histone PTM Profiles with a Yeast model and Human Blood using Quantitative Mass Spectrometry.
Hersman E, Wang A, Mitchell L, Boeke J, and Cotter R

2013 **Poster at the International Society for the Study of Xenobiotics**
Metabolism of the Anti-HIV Drug Efavirenz and Proteomic Analysis of Cytochrome P450 Expression in Murine Liver and Brain.
Hersman E, Bumpus N

2011 **Poster at the American Society for Mass Spectrometry**
Studying Age and Calorie Restriction Histone PTM Profiles with a Yeast model and Human Blood using Quantitative Mass Spectrometry.
Hersman E, Wang A, Mitchell L, Boeke J, and Cotter R

2010 **Poster at the American Society for Mass Spectrometry**

Probing the Global PTM Profile of Histones H4 and H3: Label Dependent Shot-Gun Proteomics.

Hersman E, Wang A, Boeke J, and Cotter R